

QY	971	ACAGTTCATATGGGCTACAACTTCACGTTGCTGGTCTGGAGAGATCATCTCA	1030
Db	1004	ATACTGTTGGGGTCCATACCTCCACGGCTTACCTGCTGTATGTGGAGAGCCCTTGACGCA	1063
QY	1031	TTTGTCCTGCTGCTGCTTATACCGTCGAGCCGGGCATGGCATGTATGGCTTCACAGTTC	1090
Db	1064	TTGTGGCAGTGGCATATGGGCGCCCTTCACACAGTCTATGGCCATGACAG-----	1111
QY	1091	TCCTCAGATTGGGGAGCTCAGCTTGGCATGTGATGATGCTCTACGCTGCTGCTCTCA	1150
Db	1112	-----CAGCTGCATCTGGGTTCTCTATATCGAACHCAGCTGTTGGAAGTCTGATGA	1162
QY	1151	CAGGTTGCTCCTAATATCTTAAAAATATGGAAGAACCCTCATGAGGCTTAAATATTTGATG	1210
Db	1163	CAGGTTTAAATCTTAAAGTTGCCTCTCTGGGAGACAGCCATCTGACACAGAACTGATATGATG	1222
QY	1211	ACCAAGTTTCTGGAAGTTTCTCA	1235
Db	1223	ATTCTGTTATTTGGAAGGTCCCTTAA	1247

RESULT 4

```

US-09-799-451-362
Sequence 362, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aifeng J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Dimauc, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_FL_genes Version 2.0
SEQ ID NO 362
LENGTH: 1792
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (32) .. (1354)
US-09-799-451-362

```

Query Match	14.0%;	Score 175.4;	DB 4;	Length 1792;
Best Local Similarity	49.4%;	Pred. No. 1.1e-42;		
Matches 455; Conservative	0;	Mismatches 466;	Indels 0;	Gaps 0;

QY 141 CTAATCAAGTGGCCAAAGATCTGACCGGTATGGCCGCCAATGGCTTGGGGTCTCCCACTTC 200
 Db 211 CTACCCAAgCTTCCAGAGAGGTCAGCCATGCTCTGTGGGCTTTGACTTCTCTCAAGCT 270
 QY 201 GAGTTCCGGAGACACACTGGAGAGATGTGGCTTCACTCTTCAAGCTGCGCGCTTGG 260
 Db 271 CTTCCTGCAAGCGCTTACGGCTTCAAGAGAGGTGGCTTCACTTCTCTCCCTGGCGCGCTTTGC 330

QY	261	TGTCAGTGGGCAATCTCTGTGGAGGGCTTCTTGAGCCAGTTCCCTTCGGGAAAGTGT	320
Db	331	CTGTCACTGTGCCACACTGTGTCCAGGGCTTTTCTTCCACTCTTCCACGGTGGCACATCCA	390
QY	321	CATCACTGTTCAGTAATTCGGCTGGCCACCATGAGTGTCTTGTGGTGTCAATCTCACT	380
Db	391	TGTTGGCGTGGAGAGCATGATCATAGTCTGACTTTTGTGGGGGGCGGTGCTACTCTCTT	450
QY	381	GGATGTGTCTTGGGGGAAGTCACTTGGCGGAGTTGGTGTGATGTGTGTGTGTGTGAGGT	440
Db	451	TGGTGCGGTCTGGGGGAAGACGGGCTTACCAGGTGTGTCTCAATGAGGCTCGTGGAGGT	510
QY	441	GACAGCTTTAGGCAACTGAGAGTGGTCACTAGTAATATCTTCAACACAGACTACCAT	500
Db	511	GGTGTGTTTGGATCAATAGATTTGTGTCTCTTATCTCTGGGGGTGAAAGATGCCGG	570
QY	501	GAACTAGATGCACATTTACGTGTGGACGCTAATTTTGGGCTGTCTGTGCGCTGTGCTT	560
Db	571	AGGCTTCATGACTATCCACACTTTGTGTGCTTACTTCGGGCTGTCTTGTGGCGGTTCT	630
QY	561	GCCAAAGCCTTACCCGAGGGAAGGAGATTAAGATCAGACAGCAACGATACCCACTTT	620
Db	631	GTCACAGGCCCACTGGAGAAAGAGACGCCAGGGCTTCCTTACCATTTAGACT	690
QY	621	GTCTGCATGTGGGCGCCCTCTTCTTGAGATTTCTGGCCAAATTTCACATCTGCTCT	680
Db	691	CTTGCCATGATTTGGGACCATCTTCTGTGATCTTGTGGCTTAGCTTCAATGTCTGCACT	750
QY	681	GCTGAGAAATCCAAATCGAAAGAGATGCCGTGTTCAACACTATCTATGTCTGAGACT	740
Db	751	CACAGCGCTGGGGGTGGGACAGCATCGGACCGCCCTCAACATATCACTCCCTGGCTGC	810
QY	741	CAGCGTGTGACAGCATCTCAGGGTCAATCTGTGGCTCACCCCAAGGGAAGATCAGCAA	800
Db	811	CAGCACCTTGGACACTTTCCTTGTCTGTACGCCCTTGTAGGGGAAAGTGGAGGCTTGACAT	870
QY	801	GACTTATGTGCACAGTGGCGGTGTGTGGACAGAGGCGGTGGCTGTGGGTACCTGTGTCACT	860
Db	871	GGTCCACATCCAAATATGACGCGCTGGCTGGAGGGGTGTGGTGGGACCTCAAGTGAAT	930
QY	861	GATCCCTTCTCGTGGCTTGGCCATGTGTGTGGGTCTGTGTGGTGGCTGATCTTCCGTGG	920
Db	931	GATGTGACACCTTTTGGGGCTCTGGGACGCTGGCTTCTTGGCTGGGACTGTCTCCACGCT	990
QY	921	GGGAGCCAAAGTACCTGCGGGGTGTGTGAACGAGTGTGGGGATTTCCCAACAGCTCAT	980
Db	991	GGGGTACAAAGTTCTTCAAGCCCATCTTGGAATCAAATTCAAAGTCCAAACACATGTGG	1050
QY	981	CATGGGCTACAACTTTCAGCTTGTGGGTCTGTGTGAGAGATCATCTTCAATGTGTGCT	1040
Db	1051	AGTCCAAACCTCCATGGGATGCCGGGGGTCTGTGGGGCCCTCGTGGGGTCTCTTGTGGC	1110
QY	1041	GGTGTGATATCCGTGGAGC	1061
Db	1111	TGGACTGTCAACCATGAAAGC	1131

RESULT 5

US-09-949-016-3178
 / Sequence 3178, Application US/09949016
 / Patent No. 6812339
 /
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / WITH THE INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: C1001307
 /
 / CURRENT APPLICATION NUMBER: US/09/949,016
 /
 / CURRENT FILING DATE: 2000-04-14
 /
 / PRIOR APPLICATION NUMBER: 60/241,755
 /
 / PRIOR FILING DATE: 2000-10-20
 /
 / PRIOR APPLICATION NUMBER: 60/237,768
 /
 / PRIOR FILING DATE: 2000-10-03
 /
 / PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FaSeq for Windows Version 4.0
 ; SEQ ID NO 3778
 ; LENGTH: 1939
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-3778

Query Match 13.2%; Score 165.4; DB 4; Length 1939;
 Best Local Similarity 49.1%; Pred. No. 1,3e-39;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTGGCCAGATCTGACCGTATGCGCGCCATTGGCTTGGGCTTCTCACTC 200
 DB 201 CTACCCAGGCTTCAGAGACGTCAGTGTCTTGTGGGCTTGGCTTCTCTCAAGAC 260
 QY 201 GAATTCGAGAGACACGCTGAGAGAGTGGCTTCAACCTCTTCATGCTGGCGCTTG 260
 DB 261 TTTCCTGACAGCGCTACCGCTTACGCGCGCTGGCGCTTCACTTCTGTTGGCAGCGCTT 320
 QY 261 TGTGAGTGGGCAATCCTGCTGACGCGCTTCTGAGCCATTCCTCTGGAAGGTGT 320
 DB 321 CATTCAGTGGCGCTGCTCAAGAGGCGTGGCTTCACTTCTTACAGACCGCTACATGT 380
 QY 321 CATCACACTGTTCAATTCGCGCTGGCCACCATGAGTGTCTTGGCTGATCTCAGT 380
 DB 381 CGTGGGCGTGGAGAACTCATCAACGCTGACTTCTGGTGGCGCTCTGTCTGGCGCTT 440
 QY 381 GATGCTGCTTGGGGAAGGTCAACTTGCGCAGTTGGTGTGATGCTGTGTGAGGT 440
 DB 441 TGGGGAGTTCTGGGTAAAGTCAAGCCCATTCAGCTCTCATCATGACTTCTTCCAGT 500
 QY 441 GAGAGCTTTAGGCAACCTGAGATGTCATCAGTAATATCTTCAACACACATACCAT 500
 DB 501 GACCCCTTTCGCTGTAATGATTCATTCCTTAACCTCTAAAGGTGAAGATGCAAG 560
 QY 501 GAACATGATGACATCTACGCTGTTCGAGCCTATTTTGGGCTGCTGTGGCTGTGCT 560
 DB 561 AGGCTCATATGACATCCACACATTTGGCGCTACTTTGGGCTCAGAGTACCGGATCT 620
 QY 561 GCCAAGGCTCTAACCGGAGAAAGAGATTAAGATCAGACAGCAACGATCCCAAGTT 620
 DB 621 CTACCGACGCAACTAGAGCAGAGCAAGAGAGACAGAAATTCGTGTAACAGTGGACT 680
 QY 621 GTTGCATGCTGGGCGCCTCTTCTGTGATGTTCTGGCCAAAGTTTCAACTCTGCT 680
 DB 681 CTTTGCATATGATGGACCCCTCTTCTGTGATGTAATGCTGGCCCAAGCTTCAACGCA 740
 QY 681 GCTGAGAGTCCATCGAAGAAAGAAATGCGGTGTTCAACACCTACTATCTGTAGCAGT 740
 DB 741 ATCTCTACATGGGAGACGACGACGAGCGCCCATATCAACTATCTGCTCTTGGCAGC 800
 QY 741 CAGCGTGTGACAGCATCTCAGAGGTATCTTGGCTCACCCCAAGGAAATCAGCAA 800
 DB 801 CTGCGTCTTACTCGGTGGCAATATCAGTGCCTCTGCAAGAAAGGCAAGCTGAGCAT 860
 QY 801 GATTTATGTGCAAGTGGCGGTGTTGGAGAGAGCGCTGTGGGTATCTCTGTCACT 860
 DB 861 GGTGCAATCAAAATGTCACGCTCGAGAGAGGAGGTGGCGGTGTAACCTCTGTGAT 920
 QY 861 GATCCCTTCCGCTGGGCTGCAATGCTGTGGGCTTGTGGGCTGATCTCCGCGG 920
 DB 921 GATGCTCATGCTTACGCTGCTCTCATCATCGGCTTGTGTGGGCTATCTTCCACCT 980
 QY 921 GGAAGCCAGTACCTGCGGAGGTGTGTAACGAGTGTGGGATTTCCCAAGCTTCAT 980
 DB 981 GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAATCCAGGACATATGTG 1040
 QY 981 CATGGGTTAACAATCTTGTGCTGGGTGTGCTTGGAGATCATCTACATTTGTG 1035
 DB 1041 CATTTAACAATCTTGTGCTGGGTGTGCTTGGAGATCATGAGGAGCATGTGGGTGTG 1095

RESULT 6
 US-09-461-325-75
 ; Sequence 75; Application US/09461325A
 ; Patent No. 6475753
 ; GENERAL INFORMATION:

; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 94 Human Secreted Proteins
 ; FILE REFERENCE: P2029P1
 ; CURRENT APPLICATION NUMBER: US/09/461,325A
 ; CURRENT FILING DATE: 1999-12-14

; EARLIER APPLICATION NUMBER: PCT/US99/13418
 ; EARLIER FILING DATE: 1999-06-15

; EARLIER APPLICATION NUMBER: 60/089,507
 ; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,508
 ; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,509
 ; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,510
 ; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/090,112
 ; EARLIER FILING DATE: 1998-06-22

; EARLIER APPLICATION NUMBER: 60/090,113
 ; EARLIER FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 75
 ; LENGTH: 1650

; TYPE: DNA
 ; ORGANISM: Homo sapiens

; US-09-461-325-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
 Best Local Similarity 48.7%; Pred. No. 1.2e-22;
 Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

QY 364 TCGGCTGATCTCAGTGAATGCTGTCTGGGAGAGTCAACTTGGCGAGTGTGTG 423
 DB 33 TCTGTCTGCTGGCCTTTGGGCGAGTCTGGTGAAGTCAAGCCCATTCAGCTGCTATC 92
 QY 424 ATGTCTCGTGGAGGTGACAGCTTTAGGCAACCTGAGATGCTCATCTAGTAATATCTTC 483
 DB 93 ATGACTTTCTTCAAGATGAGCCCTCTTCGCTGATGATGATTCATTCCTTAACCTGCTA 152
 QY 484 AACAGACTACCAACATGAACATGATGACATCTTACGTGTTCAGCCTTATTTGGGCTG 543
 DB 153 AAGGTAAAGATGACAGAGGCTTCATGACCATTCACACATTTGGCGCTTATTTGGGCTC 212
 QY 544 TCTGTGGCCTGAGGCTGCAAGGCTCTTACCAGAGGAAAGGAGTAAAGATCAGACA 603
 DB 213 ACAGTAAACCGGATCTCTTACAGACCAACTTATGAGACAGCAAGAGAGACAGAAATCT 272
 QY 604 GCAACGATACCCAGTTGTCTGTCATGCTGGGCGCCTCTTCTTGTGATGTTCTGGCCA 663
 DB 273 GTGTACAGTGGAGCCTCTTGTGCAATGATGAGCAACCTCTTCTGTGATGATCTGGCCC 332
 QY 664 AGTTTCACTGTCTCTGCTGAGAGTCCATGAGAAAGAAATGCGGTGTTCAACACC 723
 DB 333 AGCTTCAACTCAGCCATATCTTACATGAGGAGACAGCAGACCGGCAATCAACACC 392
 QY 724 TACTATGCTGTAGCAGTACGCGGTGAGACAGCATCTCAGGATCATCTTGGCTACCCC 783
 DB 393 TACTGCTCTTGGCAGCCTGCGTGTACTTACTCGGTGAGCAATATC-CAGTCCCTGCAAG 451
 QY 784 CAAGGAGATCAGCAAGACTTATGTGCAAGTGCAGTGTGTGGCAGAGGCGTGTGTG 843
 DB 452 AAGGGCAAGGTGACATGTGAGCAATCCAGAAATGCAACGCTGCAAGAGGAGGTGGCGGTG 511
 QY 844 GGTACTGCTGTGTAATCTTCTTCTGCTGGCTTGCATGATGCTGTGTGTGCT 903
 DB 512 GGTACGCTGTGAGATGATGCTCATGCTTACGAGTCCCTTCACTCATGCTGTGTGCTG 571

```

QY 904 GGGCTGATCTCCGCGGGAGGCCAGTACCTGCGGGGTTGTGTAACGAGTGGCTGGG 963
DB 572 GGCAATCTCCACCTGGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAC 631
QY 964 ATTCCCAAGCTCCATCATGAGGCTACAACTTCACTTGTGGGTCTGCTTGGAGATC 1023
DB 632 ATCCAGAGACATGTGGCATTTAACTGTGATGGCAATCTTGGCATCATAGGGGATC 691
QY 1024 ATCTACATTGTG 1035
DB 692 GTGGTGTCTGTG 703

```

```

RESULT 7
US-10-012-542-75
; Sequence 75, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-75

```

```

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
Best Local Similarity 48.7%; Pred. No. 1.2e-22;
Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

```

```

QY 364 TCGGTGCTGATCTCAAGTGAATGCTGCTTGGGGAAGTCACTTGGCGCAGTTGGTGTG 423
DB 33 TCTGTGCGCTGGCCTTTGGGGCAGTTCTGGTAAGTCAAGCCCATTCAGCTGCTCATC 92
QY 424 ATGTGCTGTGAGGTGAAGTGAAGCTTTAGGCACTTGAGATGCTCATCAATATCTTC 483
DB 93 ATGACTTTCTTCCAAAGTGAACCTCTTGTGTAATGAGTTCAATTCCTTAACTGCTA 152
QY 484 AACACAGACTACCACTGAATGAATGATGACATCTACGTTGTGGCAGCTATTTTGGGCTG 543
DB 153 AAGGTGAAGATCAGAGAGGCTCTCATGACATCAACATTTGGCGCTTACTTTGGGCTC 212
QY 544 TCTGTGAGCTGTGCTGCCAAAGCCTCTTACCCGAGGGAACGAGAGATTAAGATCAGAC 603
DB 213 ACAGTAGCCGGAGTCTTACCGACGCACTTAGAGAGAGCAAGGAGACGAATTTCT 272
QY 604 GCAACATATCCAGTTTGTCTGCCATGCTGGCGCCCTCTTCTTGTGATGTTCTGGCCA 663
DB 273 GTGTACAGTGGACCTTTTGGCATGATTTGGCACCTCTTCTGTGATGTACTGGCCC 332
QY 664 AGTTTCAACTCTGCTGCTGCTGAGAGTCCAAATGGAAGAAAGAAATGCCGTTTCAACACC 723
DB 333 AGCTTCAACTCAACCATATCTTCAATGGGAGACGACGACCGAGCCGCATCAACACC 392

```

```

QY 724 TACTATGCTGTAGACATGACGCTGTGAGACCATCTCAAGGTCATCTTGGCTCACCC 783
DB 393 TACTGTCTTGGACACCTGCGTGTCTTACCTGATGGAATATC-CAGTGCCCTGCAAG 451
QY 784 CAAGGGAAGATCAGCAAGACTTATGTGCAAGTGGCGGTGTGGCAGAGGCGTGGTG 843
DB 452 AAGGGCAAGGTGACATGTGTGCAATCCAGAAATGCCACCTCCGAGAGGGGGCGTG 511
QY 844 GGTACCTCGTGTACCTGATCCCTTCTCCGTGGCTTGGCATGGTGTGCTGTGGCT 903
DB 512 GTTACCTGCTGTGAGATGATGCTCATGCTTACGTTGCTCATTCAGCTTGTCTG 571
QY 904 GGGCTGATCTCCGCGGGAGGCCAGTACTCTGCGGGGTGTGTAAACGAGTCTGGG 963
DB 572 GGCAATCTCCACCTGGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAC 631
QY 964 ATTCCCAAGCTCCATCATGAGGCTACAACTTCACTTGTGGGTCTGCTTGGAGATC 1023
DB 632 ATCCAGAGACATGTGGCATTTAACTGTGATGGCAATCTGATGCTTGGCATCATAGGGGATC 691
QY 1024 ATCTACATTGTG 1035
DB 692 GTGGTGTCTGTG 703

```

```

RESULT 8
US-10-115-123-75
; Sequence 75, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P202930AP12
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-75

```

```

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
Best Local Similarity 48.7%; Pred. No. 1.2e-22;
Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

```

```

QY 364 TCGGTGCTGATCTCAAGTGAATGCTGCTTGGGGAAGTCACTTGGCGCAGTTGGTGTG 423
DB 33 TCTGTGCTGTGCGCTTTGGGCAATTTCTGGTAAGTCAAGCCCATTCAGCTGCTCATC 92
QY 424 ATGTGCTGTGAGGTGAAGTGAAGCTTTAGGCACTTGAGATGCTCATCAATATCTTC 483
DB 93 ATGACTTTCTTCCAAAGTGAACCTCTTGTGTAATGAGTTCAATTCCTTAACTGCTA 152
QY 484 AACACAGACTACCACTGAATGAATGATGACATCTACGTTGTGGCAGCTATTTTGGGCTG 543
DB 153 AAGGTGAAGATCAGAGAGGCTCTCATGACATCAACATTTTGGCGCTTACTTTGGGCTC 212

```

Qy	544	TCGTGGCCGTGGGCTCGCCCAAGGCTCTACCCGAGGGGAACGGAGATTAAGATCAACA	603
Db	213	ACAGTAAACCCGATCCTCTTACCGAGGCAACCTAGAGCAGACAAAGACAGAAATTC	272
Qy	604	GCAACGATACCCAGTTTGTCTGCATGCTGGGCGCCCTCTTCTTGTGTGATGTTCTGGCCA	663
Db	273	GTTACACAGTCGGACCTCTTGTCCATGATGGACCTCTTCTGTGTGATGTACTGGCCC	332
Qy	664	AGTTCAACCTGTCTGTCTGTGAGAAAGTCCAAATCGAAAGAAAGATGCCGTGTCAACCC	723
Db	333	AGTTCAACCTCAACCATATCTTACCATGAGGAGACGACACCGAGCCCGCATTAACCC	392
Qy	724	TACTATGCTGTAGCACTAGCGGTGTGACAGCCATCTCAGGGTCATCTTGGCTCACCCC	783
Db	393	TACTGTCTTGTGGCAGCTGCGTGTACTCTCGGTGGCAATATC-CAGTGCCTCGCAAG	451
Qy	784	CAGGGAAGATCAGCAAGACTTATGTGCACATGGCGGTGGCAGAGAGGTGGCTGTG	843
Db	452	AAGGGCAAGGTGACATGTGTGCACATCCAAATGACACGCTGGCAGAGAGGGGTGGCTG	511
Qy	844	GGTACCTCGTGTCACTGATCCCTTCTCCGTGGCTGTGCATGTAGTCTGGGTCTTGTGGCT	903
Db	512	GGTACCGGTGTGAGATGATGTCTCATGCCCTTACGGGTGCTCATCATGCTGTGTGTC	571
Qy	904	GGGCTGATCTCCGTGGGGGGAGCCAAAGTACTTCCGGGGGTGTGTAAACCAAGTCTGGGG	963
Db	572	GGGATATCTCAACCTCGGGTTTTGTATACCTGACCCCAATCTCTGAGATCCCGGGCTGAC	631
Qy	964	ATTCCCCACAGCTCCATCGAGGCTTACAATTCACCTTGTGGGTCTGTCTGTGAGATCT	1021
Db	632	ATCCAGAGACACATGTGGCATTAAACAATCTGATGCAATTCCTGGGATCATATGGGGGATC	691
Qy	1024	ATCTCAATTGTG 1035	
Db	692	GTGGGTGCTGTG 703	

RESULT 9
 US-09-949-016-12052/c
 Sequence 12052, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: C1001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 12052
 LENGTH: 34230
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-12052

	Query Match	4.5%	Score 56.6;	DB 4	Length 34230;
	Best Local Similarity	52.3%	Pred. NO.1.2e-05;		
	Matches	125; Conservative	0; Mismatches 11;	Indels	Gaps 0;
Oy	229	GTGGCTTCACACTTTCATGCTGGCGGTTGGTGAGTGAGGGGCATCCTGTGAACGC	288		
Dd	4663	GTGGCGGTGATGCGTGTATCATGTGATGTGGTGGCTGGTGATGTGATGTGATGTG	4604		
Oy	289	TTCCTGAGCCACTTCCCTTTCTGGAGAAGGTGTCAACACTTTCAGATTTCGGCTGCC	348		
Dd	4603	GTGGTGATGATGCGTGGTGGTGGTGGTGGTGATGTGGATGGTGGTGGTGGTGGTGG	4544		

[illegible]

```

RESULT 10
US-09-949-016-13765/c
: Sequence 13765, Application US/09949016
: Patent No. 681239
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13765
: LENGTH: 128470
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-13765

```

[illegible]

```

RESULT 11
US-09-949-016-49987/c
; Sequence 49987, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49987
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-49987

Query Match 4.5%; Score 56.2; DB 4; Length 601;
Best Local Similarity 56.1%; Pred. No. 1.3e-06;
Matches 106; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 613 CCCAGTTGTTCTGCCATCTGGGCGCCCTCTTTTGTGATGTTCTGGCAAGTTTCAAC 672
DB 601 CCCCTCTTTGTTGTTCACTACAGAGGAGCTCTTTCTGTGATGTTTGGCCAGCTTTAAC 542
QY 673 TCTGCTGTGCTGAGAGTCCATCGAAGAAAGATGCCGTTTCAACACTACTATGCT 732
DB 541 TCGGCATTGCTGAACCTGAGACAAACAGTGCAGGGGCATTGTAACACGTACTTCTCT 482
QY 733 GTAGACAGTCAGCGTGTGACAGCAGCATCTCAGGTCATCTTGGCTCACCCCAAGGAG 792
DB 481 CTCGCTGCTGTGTGTCTCAGACGCTTTCCTTCCAGCTTAGTGAGACACCGAGCGAAG 422
QY 793 ATCAGCAAG 801
DB 421 CTCACACATG 413

RESULT 12

US-09-949-016-13199
Sequence 13199, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13199
LENGTH: 35676
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(35676)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13199

Query Match 4.5%; Score 56.2; DB 4; Length 35676;
Best Local Similarity 56.1%; Pred. No. 1.7e-05;
Matches 106; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 613 CCCAGTTGTTCTGCCATCTGGGCGCCCTCTTTTGTGATGTTCTGGCAAGTTTCAAC 672
DB 23970 CCCCTCTTTGTTGTTCACTACAGAGGAGCTCTTTCTGTGATGTTTGGCCAGCTTTAAC 24029
QY 673 TCTGCTGTGCTGAGAGTCCATCGAAGAAAGATGCCGTTTCAACACTACTATGCT 732
DB 24030 TCGGCATTGCTGAACCTGAGACAAACAGTGCAGGGGCATTGTAACACGTACTTCTCT 24089
QY 733 GTAGACAGTCAGCGTGTGACAGCAGCATCTCAGGTCATCTTGGCTCACCCCAAGGAG 792
DB 24090 CTCGCTGCTGTGTGTCTCAGACGCTTTCCTTCCAGCTTAGTGAGACACCGAGCGAAG 24149

QY 793 ATCAGCAAG 801
DB 24150 CTCACACATG 24158

RESULT 13

US-09-949-016-15520
Sequence 15520, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15520
LENGTH: 29172
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(29172)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15520

Query Match 4.3%; Score 53.8; DB 4; Length 29172;
Best Local Similarity 62.0%; Pred. No. 8e-05;
Matches 85; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 153 CCAGATCTGACCGGATGAGCGGCGCATTTGGGCTTCTCACTGAGATTCCGAG 212
DB 11588 CAGAGCGTGACGATGATGCTTGTGGGCTTGGCTTCTCACTGAGATTCCGAGCG 11647
QY 213 ACACAGCTGAGCAGTGTGGCCTTCAACCTCTTCATGCTGGGCGTGTGACAGTGGAG 272
DB 11648 CTACGGCTTCAGCGCGCGGTTCACACTCTGTGTGACAGCTTCGCGATCCAGTGGAG 11707
QY 273 AATCTGCTGACGGCT 289
DB 11708 GCTGCTCATGACAGGCT 11724

RESULT 14

US-09-949-016-49982/c
Sequence 49982, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49982
LENGTH: 601

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 23:33:49 ; Search time 5459 Seconds
(without alignments)
1130.764 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1 atgagctcctagaccgcg.....attggctgtgatttcaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_on:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_str:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	6	AX022514 Sequence
2	1254	100.0	1254	6	BD124102 Novel nuc
3	1252.4	99.9	1254	9	AB018969 Homo sapi
4	1252.4	99.9	1354	9	AB018969 Homo sapi
5	1252.4	99.9	1545	9	HSRHIIT
6	1252.4	99.9	2790	9	HSRHIIT
7	1250.8	99.7	1254	9	AB018966 Homo sapi
8	1250.8	99.7	1254	9	AB018967 Homo sapi
9	1250.8	99.7	1480	9	AF510069 Homo sapi
10	1250.8	99.7	1480	9	AF510070 Homo sapi
11	1247.8	99.5	1251	9	AJ784309 Homo sapi
12	1247.8	99.5	1251	9	AJ784310 Homo sapi
13	1247.8	99.5	1251	9	AJ784311 Homo sapi
14	1247.8	99.5	1251	9	AY449381 Homo sapi
15	1247.8	99.5	1251	9	AY449382 Homo sapi
16	1247.8	99.5	1251	9	AY449383 Homo sapi
17	1247.8	99.5	1251	9	AY449384 Homo sapi
18	1247.8	99.5	1251	9	AY449385 Homo sapi
19	1247.6	99.5	1260	9	S78509 Rhd [human,

20	1246.2	99.4	1251	9	AJ704215 Homo sapi
21	1246	99.4	1254	9	AB018968 Homo sapi
22	1246	99.4	1254	9	S82449 Homo sapi
23	1242.8	99.1	1458	9	AF187846 Homo sapi
24	1239.8	98.9	1251	9	S57971 Rh polypept
25	1239.6	98.9	1254	9	AB046420 Homo sapi
26	1238	98.7	1328	9	AB049754 Homo sapi
27	1233.4	98.4	1248	9	AJ784312 Homo sapi
28	1231.4	98.2	1247	9	AF037626 Homo sapi
29	1226.8	97.8	1254	9	AB037270 Homo sapi
30	1225.2	97.7	1336	9	AF132679 Homo sapi
31	1222	97.4	1330	9	HS257026 Homo sapi
32	1220.4	97.3	1254	9	AF177938 Homo sapi
33	1199.6	95.7	1328	9	AB030388 Homo sapi
34	1199.6	95.7	1328	9	AB049753 Homo sapi
35	1193.2	95.2	1303	9	BC075081 Homo sapi
36	1191.6	95.0	1297	9	AY603478 Homo sapi
37	1191.6	95.0	1350	9	GORRHB
38	1190	94.9	1430	9	AF510065 Homo sapi
39	1188.4	94.8	1328	9	AB018644 Homo sapi
40	1186.8	94.6	1350	9	CHPRHC
41	1186.8	94.6	1430	9	AF510067 Homo sapi
42	1186.8	94.6	1430	9	AF510068 Homo sapi
43	1185.2	94.5	1384	6	I89819 Sequence 1
44	1185.2	94.5	1384	6	I89819 Sequence 1
45	1185.2	94.5	1471	9	HSRHI30A Human mRNA

ALIGNMENTS

RESULT 1
AX022514 LOCUS 1254 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 41 from Patent WO9937763.
ACCESSION AX022514
VERSION AX022514.1 GI:10046112
KEYWORDS
SOURCE .
ORGANISM unidentified
unidentified
unclassified.
REFERENCE 1
AUTHORS Flegel,W.A. and Wagner,F.F.
TITLE Novel nucleic acid molecules correlated with the rhesus weak d phenotype
JOURNAL Patent: WO 9937763-A 41 29-JUL-1999:
ELEGEL WILLY A (DE) ; WAGNER FRANZ F (DE) ; DRK BLUTSPENDEDIENST
BADEN-WUE-TBE)

FEATURES

source
1..1254
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..1254
/note="unamed protein product"
/codon_start=1
/protein_id="CAC07879.1"
/db_xref="GI:10046113"
/translation="MSSKTPRSVRCPLMALTLBAALLFFFTTHYDNLSEKGL
VASVYQVGDITVAAIGLGLTSSFRHSWSVAFLFMALGVQWILLDGLISQRP
SGKVVITLFSIRLATWSALSVLISVDAVLQKVMIAQLVVWLVETVLMGLRVYISNI
FNTDYNNMMHIVFAVFGLSVAMCLPKLPETGEDKQDQATLPSLAMIAGALFLNM
FWPSPNSALRSPRIERKNVFNTRYVAVVAVVTRAISSSLAHQGRKISKTYHSAVLA
GVAVGTSGHLRSPWLAIVLWGLISVGAICVLPCCNRVLTGSHSIMGVNFSL
LGLGERTIVTLVLDITVAGNGMIGFQVLSTIGELSLAIVIALTSLTGLLNLKTI
WKAHPRAKYFDQVFWKPHLAVGP"

ORIGIN

Query Match 100.0%; Score 1254; DB 6; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 ATGAGCTTAAGTACCGCGGCTGTGTCGGCGCTGCTGCGCCCTCTGGGCGCTAACATCG 60

```

Db      1 ATAGGCTTAAGAACCGCGGCTGTCGCGCGCTGCTCCCTCTGGGCGCTTAACACTG 60
Qy      61 GAAGAGCTCTCATTTCTCCCTCTCTATTTTATTTTAACTTAAGACGCTTCTTAAAGAT 120
Db      61 GAAGAGCTCTCATTTCTCCCTCTCTATTTTATTTTAACTTAAGACGCTTCTTAAAGAT 120
Qy      121 CAAAAGGGGCTCGTGGCATCTTAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT 180
Db      121 CAAAAGGGGCTCGTGGCATCTTAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT 180
Qy      181 GCGTTGGGCTTCTCATCTCGAGTTCCGGAACAACAGCTGGAGAGTGTGCTTCAAC 240
Db      181 GCGTTGGGCTTCTCATCTCGAGTTCCGGAACAACAGCTGGAGAGTGTGCTTCAAC 240
Qy      241 CTCTTCATGCTGCGCTGTGTGTCAGAGTGGCAATCTGCTGGAGGCTTCTTGAGCCAG 300
Db      241 CTCTTCATGCTGCGCTGTGTGTCAGAGTGGCAATCTGCTGGAGGCTTCTTGAGCCAG 300
Qy      301 TTCCCTTCTGGGAAGGTGTATCAACTGTTCACTGATTCGCTGGCCACCATGAGTGTCT 360
Db      301 TTCCCTTCTGGGAAGGTGTATCAACTGTTCACTGATTCGCTGGCCACCATGAGTGTCT 360
Qy      361 TTGTCGTCGTCATCTCATGTGATCTGTTGGGGAAGCTCACTTGGCCAGTGTGTG 420
Db      361 TTGTCGTCGTCATCTCATGTGATCTGTTGGGGAAGCTCACTTGGCCAGTGTGTG 420
Qy      421 GTGATGTGTGTGTGAGAGTGAACAGCTTTAGGCAACCTGAGATGTATCATGATATATC 480
Db      421 GTGATGTGTGTGTGAGAGTGAACAGCTTTAGGCAACCTGAGATGTATCATGATATATC 480
Qy      481 TTCAAACAAGACCTCAATATGAATGACATCTACGTGTGTGCAAGCTTATTTTGG 540
Db      481 TTCAAACAAGACCTCAATATGAATGACATCTACGTGTGTGCAAGCTTATTTTGG 540
Qy      541 CTGTCTGTGTCGTGTGCTGCTGCAAGAGCTCTACCCGAGGGAACGAGATTAAGATCAG 600
Db      541 CTGTCTGTGTCGTGTGCTGCTGCAAGAGCTCTACCCGAGGGAACGAGATTAAGATCAG 600
Qy      601 ACAGCAACGATACCAAGTTTGTCTGCAATGTGTGGGCGCTCTTCTTGTGTGATTTCTG 660
Db      601 ACAGCAACGATACCAAGTTTGTCTGCAATGTGTGGGCGCTCTTCTTGTGTGATTTCTG 660
Qy      661 CCAAGTTTCAACTGTGCTGTGAGAGTCCAATGCAAGGAAGATGCGTGTTCAC 720
Db      661 CCAAGTTTCAACTGTGCTGTGAGAGTCCAATGCAAGGAAGATGCGTGTTCAC 720
Qy      721 ACCTACTATGCTGTAGAGTCAAGGTGTGACAGCCATCTCAAGGTCTATCTTGGCTCAC 780
Db      721 ACCTACTATGCTGTAGAGTCAAGGTGTGACAGCCATCTCAAGGTCTATCTTGGCTCAC 780
Qy      781 CCCCAAGGGAAGATCAGCAAGACTTATATGTGCAACAGTCCGTTGGCAAGGCGTGGCT 840
Db      781 CCCCAAGGGAAGATCAGCAAGACTTATATGTGCAACAGTCCGTTGGCAAGGCGTGGCT 840
Qy      841 GTGGGTAACCTCGTGTCACTGATCCCTTCCGTGGCTTGCCATGTGTGCTGTGG 900
Db      841 GTGGGTAACCTCGTGTCACTGATCCCTTCCGTGGCTTGCCATGTGTGCTGTGG 900
Qy      901 GCTGGGCTGATCTCCGTCGGGGAGGCAAGATCTGCGGGGTGTGTAACCAAGTGTG 960
Db      901 GCTGGGCTGATCTCCGTCGGGGAGGCAAGATCTGCGGGGTGTGTAACCAAGTGTG 960
Qy      961 GGGATTCCTCCCAAGCTTCAATGAGCTTCAAGTTTCTGAGTGTGCTTGAAGAG 1020
Db      961 GGGATTCCTCCCAAGCTTCAATGAGCTTCAAGTTTCTGAGTGTGCTTGAAGAG 1020
Qy      1021 ATCATTTAATGTTGTGTGTGCTTGAATACCGTCGAGCGGCAATGATGATGAGC 1080
Db      1021 ATCATTTAATGTTGTGTGTGCTTGAATACCGTCGAGCGGCAATGATGATGAGC 1080
Qy      1081 TTCGAGGTCCTGCTGATGAGATTGGGAATCAGCTTGGCATGTATAGCTTCAAGTCT 1140

```

```

Db      1081 TTCAGGTCCTCCTCAGCAATGGGGAACCTCAGCTTGGCCATGATGACTTCAAGTCT 1140
Qy      1141 GGTCTCTGACAGAGTTGTGCTCTTAATCTTAATAATGGAAGACACTGTAGGCTTAA 1200
Db      1141 GGTCTCTGACAGAGTTGTGCTCTTAATCTTAATAATGGAAGACACTGTAGGCTTAA 1200
Qy      1201 TATTTTATGACCAAGTTTCTGGAATTTCTCATTTTGGCTGTGATTTTAA 1254
Db      1201 TATTTTATGACCAAGTTTCTGGAATTTCTCATTTTGGCTGTGATTTTAA 1254

RESULT 2
BD124102
LOCUS      BD124102          1254 bp      DNA          linear      PAT 18-SEP-2002
DEFINITION Novel nucleic acid molecule correlating to Rhesus weak D phenotype.
ACCESSION  BD124102
VERSION    BD124102.1  GI:23219047
KEYWORDS   JP 2002500884-A/41.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 1254)
AUTHORS   Fregel,V.A. and Wagner,F.F.
TITLE     Novel nucleic acid molecule correlating to Rhesus weak D phenotype
JOURNAL    Patent: JP 2002500884-A 41 15-JAN-2002;
           DRK BLJSPENDDIENST BADEN WUERTTEMBERG GEMBIH
COMMENT    UNidentified
           PN  JP 2002500884-A/41
           PD  15-JAN-2002
           PF  18-DEC-1998  JP 2000528671
           PR  23-JAN-1998  EP  98101203.2
           PI  VILLY A FREGEL,FRANZ F WAGNER
           PC  C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
           10,
           PC  C12P21/02,C12P21/08,C12Q1/02,C12Q1/68,G01N33/566,C12N15/00,PC
           C12N5/00
           CC  Strandedness: Single;
           CC  Topology: linear;
           CC  Novel nucleic acid molecule correlating to Rhesus weak D
           FH  Key phenotype
           FT  CDS
           source      Location/Qualifiers
                   1..1254
                   /organism="unidentified"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 1254; DB 6; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAGCTTAAGAACCGCGGCTGTCGCGCGCTGCTCCCTCTGGGCGCTTAACACTG 60
Db      1 ATGAGCTTAAGAACCGCGGCTGTCGCGCGCTGCTCCCTCTGGGCGCTTAACACTG 60
Qy      61 GAAGAGCTCTCATTTCTCCCTCTCTATTTTATTTTAACTTAAGACGCTTCTTAAAGAT 120
Db      61 GAAGAGCTCTCATTTCTCCCTCTCTATTTTATTTTAACTTAAGACGCTTCTTAAAGAT 120
Qy      121 CAAAAGGGGCTCGTGGCATCTTAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT 180
Db      121 CAAAAGGGGCTCGTGGCATCTTAAGTTGGCCAAAGATCTGACCGTGATGGCGGCATT 180
Qy      181 GCGTTGGGCTTCTCATCTCGAGTTCCGGAACAACAGCTGGAGAGTGTGCTTCAAC 240
Db      181 GCGTTGGGCTTCTCATCTCGAGTTCCGGAACAACAGCTGGAGAGTGTGCTTCAAC 240
Qy      241 CTCTTCATGCTGCGCTGTGTGTCAGAGTGGCAATCTGCTGGAGGCTTCTTGAGCCAG 300
Db      241 CTCTTCATGCTGCGCTGTGTGTCAGAGTGGCAATCTGCTGGAGGCTTCTTGAGCCAG 300

```

QY 301 TTCCTTCTGGGAAGGTGGTCAACAGTGTTCAGATTTGGGCTGGCCACATGAGTGTCT 360
 DB 301 TTCCTTCTGGGAAGGTGGTCAACAGTGTTCAGATTTGGGCTGGCCACATGAGTGTCT 360
 QY 361 TTGTGGTGTGATTCATGAGTGTCTGTCTTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
 DB 361 TTGTGGTGTGATTCATGAGTGTCTGTCTTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
 QY 421 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACTTGAGATGGTCACTAATATC 480
 DB 421 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACTTGAGATGGTCACTAATATC 480
 QY 481 TTCAACACAGACTACACATGAATGATGACATCTACGTTGTCGAGCCTATTTTGGG 540
 DB 481 TTCAACACAGACTACACATGAATGATGACATCTACGTTGTCGAGCCTATTTTGGG 540
 QY 541 CTGTCTGTGGCTGTGGTGTGCTGCCAAGCCTCTACCCGAGGGAACGAGATAAAGATCAG 600
 DB 541 CTGTCTGTGGCTGTGGTGTGCTGCCAAGCCTCTACCCGAGGGAACGAGATAAAGATCAG 600
 QY 601 ACAGCAACGATACCCAGTTTGTGTCGATGCTGGGCGCCTCTTGTGTGATGTTCTGG 660
 DB 601 ACAGCAACGATACCCAGTTTGTGTCGATGCTGGGCGCCTCTTGTGTGATGTTCTGG 660
 QY 661 CCAAGTTTCAACTCTGCTGTGAGAACTCAATCGAAGGAAGATGCCGTGTTCAAC 720
 DB 661 CCAAGTTTCAACTCTGCTGTGAGAACTCAATCGAAGGAAGATGCCGTGTTCAAC 720
 QY 721 ACCTACTATGCTGTAGAGATGACAGCTGTGTGACACCATCTCAGGCTCATCTTGGCTCAC 780
 DB 721 ACCTACTATGCTGTAGAGATGACAGCTGTGTGACACCATCTCAGGCTCATCTTGGCTCAC 780
 QY 781 CCCCAAGGGAAGATCGAAGACTTAATGTGACAGTGGGCTGTGGAGGAGCGTGGCT 840
 DB 781 CCCCAAGGGAAGATCGAAGACTTAATGTGACAGTGGGCTGTGGAGGAGCGTGGCT 840
 QY 841 GTGGGTAACCTCGTGTCACTGATCCCTTCTCCGTGGCTTTCAGATGGTGTGGTCTTTGTG 900
 DB 841 GTGGGTAACCTCGTGTCACTGATCCCTTCTCCGTGGCTTTCAGATGGTGTGGTCTTTGTG 900
 QY 901 GCTGGGCTGATCTCCGTGGGGAAGCCAGTACCTGCGGGGCTGTGTAACCAAGTGTCTG 960
 DB 901 GCTGGGCTGATCTCCGTGGGGAAGCCAGTACCTGCGGGGCTGTGTAACCAAGTGTCTG 960
 QY 961 GGGATTCCTCCAGAGCTTCATGAGGCTTACACTTTCAGCTTGTGGCTTGGAGAG 1020
 DB 961 GGGATTCCTCCAGAGCTTCATGAGGCTTACACTTTCAGCTTGTGGCTTGGAGAG 1020
 QY 1021 ATCATCTACATTTGTGCTGTGCTGTGATACCGTGGAGCCGGCAATGGCATGATTTGGC 1080
 DB 1021 ATCATCTACATTTGTGCTGTGCTGTGATACCGTGGAGCCGGCAATGGCATGATTTGGC 1080
 QY 1081 TTCCAGGCTCTCTCCTCAGACTTGGGGAACCTCAGCTTGGCCATCGTATGAGCTTCAAGCTCT 1140
 DB 1081 TTCCAGGCTCTCTCCTCAGACTTGGGGAACCTCAGCTTGGCCATCGTATGAGCTTCAAGCTCT 1140
 QY 1141 GGTCTCTGTCAGAGTTTGTCTTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 DB 1141 GGTCTCTGTCAGAGTTTGTCTTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254
 DB 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254

RESULT 3
 AB018969 1254 bp mRNA linear PRI 03-JUL-1999
 LOCUS AB018969 Homo sapiens RndyO mRNA for Rh blood group D antigen (RHD),
 DEFINITION complete cds.
 ACCESSION AB018969
 VERSION AB018969.1 GI:5360898

KEYWORDS RndyO; Rh blood group D antigen (RHD).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (sites) Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,
 Hyodo, H., Ishikawa, Y., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
 Teuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
 TITLE Polymorphisms of Rhdva in Japanese
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 1254)
 REFERENCE Uchikawa, M., Hyodo, H. and Ishikawa, Y.
 TITLE Direct Submission
 JOURNAL Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central
 Blood Center, Research; 4-1-31, Hiroco, Shibuya, Tokyo 150-0012,
 Japan (E-mail:hyodo@hla.cb.c.jrc.or.jp, Tel: +81-3-5485-6009,
 Fax: +81-3-3406-7892)

FEATURES

source
 1..1254
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /isolate="YO"
 /db_xref="taxon:9606"
 1..1254
 /gene="RhdYO"
 /gene="RhdYO"
 1..1254
 /note="Rhd variant RndyO"
 /codon_start=1
 /product="Rh blood group D antigen (RHD)"
 /protein_id="BA82159.1"
 /db_xref="GI:5360899"
 /translation="MSSKYPRSVRCCLPMLATLLEAALLLPFFTHYDASLEDOGL
 VASVOGDLTMAATGLGLTSPFRHSMSVAFULPMALGVOMAILDGLPSOP
 SGRVHTTFSIALMTASLSVLISVAVGLKTNLADLVVVVLEVALGTLRVANSI
 FNDTHMMNHIYVPAIFGLSVAKCLPRLPDEGTDXOTATIPSLSMGLPLIM
 FWPFSALRLSPITKRKNAVFNTRYVAASVVAIVAGSSILAHQKISKTYVASVLA
 GGVAVGTSHLIPSPMLAMVLGLVAGLISVGAKYLPCCNRYLGIPIHSSIMQYNSL
 LGILGRIIVLVLDVVGANGMIGFVYLLSIGRSLAIVIALTSGLLTLNLRKI
 WKAPHBAKTFDDQVFWKPFHLAVGF"

ORIGIN

Query Match 99.9%; Score 1252.4; DB 9; Length 1254;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTCTAAGTACCCCGGGTGTGTCGGCGCTGCTGCTCTGAGGCCCTTAACATG 60
 DB 1 ATGAGCTCTAAGTACCCCGGGTGTGTCGGCGCTGCTGCTCTGAGGCCCTTAACATG 60
 QY 61 GAAGCAGCTCTATCTCCCTCTTATTTTAAACCACATAGACGCTTCTTAGAGAT 120
 DB 61 GAAGCAGCTCTATCTCCCTCTTATTTTAAACCACATAGACGCTTCTTAGAGAT 120
 QY 121 CAAGAAGGGCTCTGGCATCTTATCAAGTTGGCAAGATCTGACCGTATGGCGGCATT 180
 DB 121 CAAGAAGGGCTCTGGCATCTTATCAAGTTGGCAAGATCTGACCGTATGGCGGCATT 180
 QY 181 GGTGGGGCTTCTCACTCGAATTTCCGAGACACAGCTGGAGCACTGTGGCTTTAAC 240
 DB 181 GGTGGGGCTTCTCACTCGAATTTCCGAGACACAGCTGGAGCACTGTGGCTTTAAC 240
 QY 241 GGTCTGGGCTTCTCACTCGAATTTCCGAGACACAGCTGGAGCACTGTGGCTTTAAC 300
 DB 241 GGTCTGGGCTTCTCACTCGAATTTCCGAGACACAGCTGGAGCACTGTGGCTTTAAC 300
 QY 301 TTCCCTTCTGGGAAGGTGGTCAACAGTGTTCAGATTTGGGCTGGCCACATGAGTGTCT 360
 DB 301 TTCCCTTCTGGGAAGGTGGTCAACAGTGTTCAGATTTGGGCTGGCCACATGAGTGTCT 360
 QY 361 TTGTGGTGTGATTCATGAGTGTCTGTCTTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
 DB 361 TTGTGGTGTGATTCATGAGTGTCTGTCTTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420

QY 421 GTGATGCTGCTGTGAGGTGACAGCTTATAGCACTGAGATGCTCATAGTAATATC 480
DB 421 GTGATGCTGCTGTGAGGTGACAGCTTATAGCACTGAGATGCTCATAGTAATATC 480
QY 481 TTCAACACAGACTACCAACATGAAATGATGACATCTTACGTGTGACAGCTTATTTGG 540
DB 481 TTCAACACAGACTACCAACATGAAATGATGACATCTTACGTGTGACAGCTTATTTGG 540
QY 541 CTGTCTGTGAGCTGTGCTGCTGCAAGGCTTACCCGAGGAAACGAGATTAATATCAG 600
DB 541 CTGTCTGTGAGCTGTGCTGCTGCAAGGCTTACCCGAGGAAACGAGATTAATATCAG 600
QY 601 AAGACAGACATACCCAGCTTGTCTGCAATGCTGAGGAGCCCTCTTCTGTGATGTTCTG 660
DB 601 AAGACAGACATACCCAGCTTGTCTGCAATGCTGAGGAGCCCTCTTCTGTGATGTTCTG 660
QY 661 CCAAGTTTCACTGCTGCTGCTGAGAGTCCATTCGAAAGAGAAATGCCGTGTTCAAC 720
DB 661 CCAAGTTTCACTGCTGCTGCTGAGAGTCCATTCGAAAGAGAAATGCCGTGTTCAAC 720
QY 721 ACTTACTATGCTGTAGACAGTACGCTGTGACAGGCTCTCAGGCTCATCTTGGCTCAC 780
DB 721 ACTTACTATGCTGTAGACAGTACGCTGTGACAGGCTCTCAGGCTCATCTTGGCTCAC 780
QY 781 CCCAAGGAAAGATGACGAAACATTAATGTGACAGTGGCTGTGAGAGAGGCTGAGCT 840
DB 781 CCCAAGGAAAGATGACGAAACATTAATGTGACAGTGGCTGTGAGAGAGGCTGAGCT 840
QY 841 GTGGGTACCTGCTGTGACAGTACCTGATCCCTTCTCGTGGCTGTGACAGTGGCTGTG 900
DB 841 GTGGGTACCTGCTGTGACAGTACCTGATCCCTTCTCGTGGCTGTGACAGTGGCTGTG 900
QY 901 GCTGGGCTGATCTCCGTGTGGGAGAGCAAGTACCTGCGGGGTGTGTAACCGAGTCTG 960
DB 901 GCTGGGCTGATCTCCGTGTGGGAGAGCAAGTACCTGCGGGGTGTGTAACCGAGTCTG 960
QY 961 GGAATTTCCCAAGCTCCATCATGAGGCTACCACTTCAAGCTGTGCTGCTGTGAGAG 1020
DB 961 GGAATTTCCCAAGCTCCATCATGAGGCTACCACTTCAAGCTGTGCTGCTGTGAGAG 1020
QY 1021 ATCATCTAATTTGTGCTGTGCTGTGATACCGTCCGAGCCGCAATGAGCATGATTTGG 1080
DB 1021 ATCATCTAATTTGTGCTGTGCTGTGATACCGTCCGAGCCGCAATGAGCATGATTTGG 1080
QY 1081 TTCAGGCTCCCTCAGACATTTGGGAACTCAGCTTGGCCATGCTGATAGCTCTCAGCTCT 1140
DB 1081 TTCAGGCTCCCTCAGACATTTGGGAACTCAGCTTGGCCATGCTGATAGCTCTCAGCTCT 1140
QY 1141 GGTCTCTGACAGATTTGCTCTAAATCTTAATAATGAGAAACACCTGATAGAGGTAA 1200
DB 1141 GGTCTCTGACAGATTTGCTCTAAATCTTAATAATGAGAAACACCTGATAGAGGTAA 1200
QY 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254
DB 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254

RESULT 4
HUMRHADANTI 1354 bp mRNA linear PRI 24-AUG-1993
LOCUS Human Rhd blood group antigen mRNA, complete cds.
DEFINITION L08429.1 GI:3373390
ACCESSION L08429.1
VERSION 1
KEYWORDS Rhd blood group; Rhd blood group; antigen; blood group antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1354)
Arce,M.A., Thompson,E.S., Wagner,S., Coyne,K.E., Ferdman,B.A. and
Lublin,D.M.
TITLE Molecular cloning of Rhd cDNA derived from a gene present in

JOURNAL RHD-positive, but not Rhd-negative individuals
MEDLINE Blood 82 (2), 651-655 (1993)
PUBMED 93320449
COMMENT 8329718
FEATURES
Source
1.1354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1254
/strand_name="Rhd"
/citation=[1]
/evidence=experimental
/product="Rhd blood group antigen"
/protein_id="AAA02679.1"
/db_xref="GI:3373391"
/translation="MSKTPSRVRCPLPMALTEBALILLFPFTHYDASLEDOKGL
VASVOYGDULTVMAAIGLGFLLSPRRHSWSVAPNLPMALGVOMAILDGLSOPF
SKRVITLPSRLATMSALSVLISYDAVLGKVLALVYVLAIVETALGNLRMTISNT
FMDYHMMMHIIYFAAYGSLVAMCLPKPIPEGREDODRATTLSIAMSGLPTM
FWPSPNSALRSPRIERKNAVPMYVAIVSVVTAISGSLAHPOKSKITYVSAVLA
GGAVATGSCHLIPSPMLAMVIGVAGLISVGAATLPPCCNRRVGIPIHSSIMGNFSL
LGLLEITIIYIIVLIDTVGAGNGMIGFVLLISGLSLAIVALSGLTLNLTKI
WAPRPAKFPDQVFWKPHLAVGP"
ORIGIN
Query Match 99.9%; Score 1252.4; DB 9; Length 1354;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTGTGGCCCTTAACATG 60
DB 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTGTGGCCCTTAACATG 60
QY 61 GAAGCAGCTCATATCTCTCTTCAATTTTAAACCACTATGACGCTTCTTGAAGAT 120
DB 61 GAAGCAGCTCATATCTCTCTTCAATTTTAAACCACTATGACGCTTCTTGAAGAT 120
QY 121 CAAAAGGGCTGTGGACATCTTATCAAGTGGCCCAATGACCTGATGCGGCCATT 180
DB 121 CAAAAGGGCTGTGGACATCTTATCAAGTGGCCCAATGACCTGATGCGGCCATT 180
QY 181 GAGTTGGGCTTCCACCTGAGTATTTCCGAGACACAGCTGAGACAGTGTGGCTTCAAC 240
DB 181 GAGTTGGGCTTCCACCTGAGTATTTCCGAGACACAGCTGAGACAGTGTGGCTTCAAC 240
QY 241 CTCTTCAATGCTGGCGCTGTGTGACAGTGGGCAATCTGCTGAGCGGCTTCTGAGCGAG 300
DB 241 CTCTTCAATGCTGGCGCTGTGTGACAGTGGGCAATCTGCTGAGCGGCTTCTGAGCGAG 300
QY 301 TTCCCTTCTGGAAGGTGTGATCAACATCTTCAATGATTTGGCTGGCCACATGAGTGT 360
DB 301 TTCCCTTCTGGAAGGTGTGATCAACATCTTCAATGATTTGGCTGGCCACATGAGTGT 360
QY 361 TTGTGCTGTGATCTCAGTGTGATGCTGTGGGGAAGTCAACTTGGCGAGTTGGTG 420
DB 361 TTGTGCTGTGATCTCAGTGTGATGCTGTGGGGAAGTCAACTTGGCGAGTTGGTG 420
QY 421 GTGATGCTGTGTGAGAGTGAAGCTTTAAGCACTTGAAGATGCTATCAGTAATATC 480
DB 421 GTGATGCTGTGTGAGAGTGAAGCTTTAAGCACTTGAAGATGCTATCAGTAATATC 480
QY 481 TTCAACACAGACTACCAACATGAAATGATGACATCTTACGTGTGACAGCTTATTTGG 540
DB 481 TTCAACACAGACTACCAACATGAAATGATGACATCTTACGTGTGACAGCTTATTTGG 540
QY 541 CTGTCTGTGAGCTGTGCTGCTGCAAGGCTTACCCGAGGAAACGAGATTAATATCAG 600
DB 541 CTGTCTGTGAGCTGTGCTGCTGCAAGGCTTACCCGAGGAAACGAGATTAATATCAG 600

QY 601 ACAGCAACGATCCAGCTTTGTCTCCAGTCTGGGCGCCCTCTTCTTGATGATGTTCTGG 660
 DB 601 ACAGCAACGATCCAGCTTTGTCTCCAGTCTGGGCGCCCTCTTCTTGATGATGTTCTGG 660
 QY 661 CCAGATTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGATGCGGTCTTAAC 720
 DB 661 CCAGATTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGATGCGGTCTTAAC 720
 QY 721 ACCTACTATGCTGAGAGTCAAGGCTGCTGAGAGTCAAGGCTGCTGAGAGTCAAGGCTGCTGAG 780
 DB 721 ACCTACTATGCTGAGAGTCAAGGCTGCTGAGAGTCAAGGCTGCTGAGAGTCAAGGCTGCTGAG 780
 QY 781 CCCCAAGGAGAGATCAGCAAGACTTATGTCACAGTCCGCTGCTGAGAGAGGCGTCTGCT 840
 DB 781 CCCCAAGGAGAGATCAGCAAGACTTATGTCACAGTCCGCTGCTGAGAGAGGCGTCTGCTGCT 840
 QY 841 GTGGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 841 GTGGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 GCTGGGCTGATCTCCGTCGGGGAGAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 GCTGGGCTGATCTCCGTCGGGGAGAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 GGGATTCGCCACAGCTTCATCATGAGGCTCAACATTCAGCTTGGCTGCTGCTGCTGCTGCTGCT 1020
 DB 961 GGGATTCGCCACAGCTTCATCATGAGGCTCAACATTCAGCTTGGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 ATATATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 DB 1021 ATATATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 1081 TTCAGGCTCTCCCTCAGAGCTTGGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 1081 TTCAGGCTCTCCCTCAGAGCTTGGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 1141 GGTCTCTGCAAGATTTGCTCTTAATCTTAATAATATGAAAGCACTTCAAGGCTTAA 1200
 DB 1141 GGTCTCTGCAAGATTTGCTCTTAATCTTAATAATATGAAAGCACTTCAAGGCTTAA 1200
 QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1254
 DB 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1254
 RESULT 5
 HSRH11
 LOCUS HSRH11 1545 bp mRNA linear PRI 07-JUL-1997
 DEFINITION H.sapiens mRNA for rhesus polypeptide (RhII).
 ACCESSION X63094.1 GI:36027
 VERSION X63094.1 GI:36027
 KEYWORDS Isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
 Rhesus polypeptide.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1545)
 Le Van Kim, C., Cherif-Zahar, B., Raynal, V., Mouro, I., Lopez, M.,
 Cartton, J.P. and Colin, Y.
 Multiple Rh messenger RNA isoforms are produced by alternative
 splicing
 Blood 80 (4), 1074-1078 (1992)
 JOURNAL MEDLINE
 PUBMED 9236085
 1379850
 TITLE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED 1438298

REFERENCE 3 (bases 1 to 1545)
 AUTHORS Colin, Y.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
 A.Cabanef, 75015 Paris, FRANCE
 COMMENT See also X63094-98 & M34015.
 FEATURES
 Location/Qualifiers
 1..1545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="1p35"
 /chromosome="1"
 /clone="RhII"
 /cissue_type="bone marrow"
 /cissue_1ib="human bone marrow cDNA"
 59..1312
 /codon_start=1
 /product="rhesus polypeptide RhII"
 /protein_id="CAA4808.1"
 /db_xref="GI:36028"
 /db_xref="GOA:002161"
 /db_xref="UniProt/Swiss-Prot:002161"
 /translation="MSSKTPRSVRCLPLMALTLLEALILLPFFPHYDASLEDQKL
 VASVYQGDLYWMAALGLGFLTSSFRHSMSVAFALFMALGVQNALILDGFLSQFP
 SGRVYITLPLSLATMSALSVLSVDAVGLKVLADLVVAVLVEALALRMVINSI
 FNTDYNNMMHIIYFAAYFGLSVAMCLPKPLPEGTEDKQOTATIPSLAMGLALFLMM
 FMSFNSALLRSPILERRNAVFNTYVAIVAAVTAIIGSSLAHPQKISKYVHSAVLA
 GYAVVTSCHLIPSPMLAVLGLVAGLISYGAARKYIPGCCNPLVGIIPHSIMQYNSL
 LGLVGRITIVLVLDVTGAGNMIGFQVYLSIGELSLAIVLALVLTGLTLNLKI
 WRAPEBAKYPDQVFWKFPFLAVGF"
 ORIGIN
 Query Match 99.9%; Score 1252.4; DB 9; Length 1545;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAGCTCTAAGTACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 DB 59 ATGAGCTCTAAGTACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118
 QY 61 GAAGCAGCTCTAATCTCTCTCTTCTAATTTTATTTTACCACATGACGCTTCTTAGAGAT 120
 DB 119 GAAGCAGCTCTAATCTCTCTCTCTTCTAATTTTATTTTACCACATGACGCTTCTTAGAGAT 178
 QY 121 CAAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 179 CAAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
 QY 181 GGCCTGGGCTTCTCTCACTGAGTTTCCGAGACACAGCTGAGAGTGGCGCTTCAAC 240
 DB 239 GGCCTGGGCTTCTCTCACTGAGTTTCCGAGACACAGCTGAGAGTGGCGCTTCAAC 298
 QY 241 CTCTTCAAGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 299 CTCTTCAAGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
 QY 301 TTCCCTTTGGGAAAGTGTGATCAACCTGTTAGATTCTGGTGGCACACAGAGTCT 360
 DB 359 TTCCCTTTGGGAAAGTGTGATCAACCTGTTAGATTCTGGTGGCACACAGAGTCT 418
 QY 361 TTGTCGGTGTGATCTCAAGTATGCTGCTTGGGAAAGTCAACTTGGCGAGTGGTG 420
 DB 419 TTGTCGGTGTGATCTCAAGTATGCTGCTTGGGAAAGTCAACTTGGCGAGTGGTG 478
 QY 421 GTGATGCTGTGTGAGAGTGAACCTTTAGAGCACTGAGATGCTATCAATATATC 480
 DB 479 GTGATGCTGTGTGAGAGTGAACCTTTAGAGCACTGAGATGCTATCAATATATC 538
 QY 481 TTCAACACGACTACACATGAACTGATGCAATCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 539 TTCAACACGACTACACATGAACTGATGCAATCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598

QY 541 CTGTCTGTGGCCCTGTGTGCTGCGCAAGCCTTACCCGAGGAAACGAGGATTAAGATCAG 600
 DB 599 CTGTCTGTGGCCCTGTGTGCTGCGCAAGCCTTACCCGAGGAAACGAGGATTAAGATCAG 658
 QY 601 AAGAGCAAGATACCCGATTGTTGTCTGTGCATGTGGGGCCCTCTTCTTGTGTGAATGTTCTGG 660
 DB 659 AAGAGCAAGATACCCGATTGTTGTCTGTGCATGTGGGGCCCTCTTCTTGTGTGAATGTTCTGG 718
 QY 661 CCAAGTTTCACTGTCTGTGTGAGAACTCCAAATGCAAGAAAGAAAGAACCCGTTTCAAC 720
 DB 719 CCAAGTTTCACTGTCTGTGTGAGAACTCCAAATGCAAGAAAGAAAGAACCCGTTTCAAC 778
 QY 721 ACTTACTATGTCTGTGTGAGAACTCCGATGTGAGCAAGCAATCTCAGGGTCAATCCTTGGCTCAC 780
 DB 779 ACTTACTATGTCTGTGTGAGAACTCCGATGTGAGCAAGCAATCTCAGGGTCAATCCTTGGCTCAC 838
 QY 781 CCCCAAGGAGATACAGCAAGATTAATGTGACATGTGGGTGTTGGCAGAGGCGTGGCT 840
 DB 839 CCCCAAGGAGATACAGCAAGATTAATGTGACATGTGGGTGTTGGCAGAGGCGTGGCT 898
 QY 841 GTGGGTACCTGTGTACCTGTATCCCTTCTCCGTGGTGGCAATGTGCTGGTCTTGTG 900
 DB 899 GTGGGTACCTGTGTACCTGTATCCCTTCTCCGTGGTGGCAATGTGCTGGTCTTGTG 958
 QY 901 GCTGGGCTGATCTCCGTGGGAGGAGCAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960
 DB 959 GCTGGGCTGATCTCCGTGGGAGGAGCAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 1018
 QY 961 GGGATTTCCCAAGCTCCATCATGGGCTTACCAACTTCAAGCTTGGGGTCTTGTGGAG 1020
 DB 1019 GGGATTTCCCAAGCTCCATCATGGGCTTACCAACTTCAAGCTTGGGGTCTTGTGGAG 1078
 QY 1021 ATCATCTAATTTGTGTGCTGTGTGATTAACCGTGGAGCGGCAATGGCATGATTTGGC 1080
 DB 1079 ATCATCTAATTTGTGTGCTGTGTGATTAACCGTGGAGCGGCAATGGCATGATTTGGC 1138
 QY 1081 TTCAGGTCCTCCATGATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1140
 DB 1139 TTCAGGTCCTCCATGATTTGGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1198
 QY 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAATATATGAAAGACCTCAGTGAAGCTTAA 1200
 DB 1199 GGTCTCTGACAGGTTTGTCTCTTAATCTTAATATATGAAAGACCTCAGTGAAGCTTAA 1258
 QY 1201 TATTTATATACCAAGTTTCTTGAAGTTTCTTCAATTTGCTGTGATTTTAA 1254
 DB 1259 TATTTATATACCAAGTTTCTTGAAGTTTCTTCAATTTGCTGTGATTTTAA 1312

RESULT 6
 HSRXIII 2790 bp mRNA linear PRI 07-JUL-1997
 LOCUS H.sapiens mRNA for rhesus polypeptide (RhXIII).
 DEFINITION X63097.1 GI:36046
 ACCESSION X63097.1 GI:36046
 VERSION isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
 KEYWORDS Rhesus polypeptide;
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 2790)
 Le Van Kim, C., Cherif-Zahar, B., Raynal, V., Mouro, I., Lopez, M.,
 Carton, J.P., and Colin, Y.
 Multiple Rh messenger RNA isoforms are produced by alternative
 splicing
 BLOOD 80 (4), 1074-1078 (1992)
 JOURNAL MEDLINE 92360855
 PUBMED 1379850
 REFERENCE 2
 Le van Kim, C., Mouro, I., Cherif-Zahar, B., Raynal, V., Cherrier, C.,
 Carton, J.P., and Colin, Y.
 Molecular cloning and primary structure of the human blood group

RHD polypeptide
 Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)
 MEDLINE 93066356
 PUBMED 1438298
 REFERENCE 3 (bases 1 to 2790)
 AUTHORS Colin, Y.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
 A.Cabanef, 75015 Paris, FRANCE
 COMMENT See also X63094.98 & M34015.
 FEATURES location/Qualifiers
 source 1..2790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p35"
 /clone="RhXIII"
 /issue_type="bone marrow"
 /issue_1ib="human bone marrow cDNA"
 17..1270
 /codon_start=1
 /product="Rhesus polypeptide XIII"
 /protein_id="CAA44811.1"
 /db_xref="GI:36047"
 /db_xref="GOA:002161"
 /db_xref="UniProt/Swiss-Prot:Q02161"
 /translation="MSSKYPKPSVRCLPLMALTLLEALILFYEFTHYDASLEDDKGL
 VASVYGGDILTMALIGLPLTSFRRSSMSVAENFPLALGVOMALILDFLSOP
 SKRVITLFSRLATMSALSVLISVDALGVNLAOLVVMVLVEVTALILGRVTSNI
 FNTDYMNMHLYVFAAYRGLSVAMCLPKPIPEGTEDDDQVATITPSLNLGALPLMI
 FWPSPNSALLSPTRKRAYPNTYAYAVSVTAISGSLHPGKISKYVHSAVLA
 GVAVGTCHLIPSPMLAMVGLVAGLISVGAARYLPPCCNRVIGIPIHSSIMGYNFSL
 LGLLEILYIVLVLDVAGNEMIIGPVVLLISGLSLAIVIALTSLTLTLNLKI
 WKAPHEARYFDQVFWKPHLAVGF"

CDS

Query Match 99.9%; Score 1252.4; DB 9; Length 2790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 ORIGIN
 1 ATGAGCTTAAATGACCCGCGGTGTGTCCGCGCTGCTGCCCTCTGAGCCCTTAACACTG 60
 17 ATGAGCTTAAATGACCCGCGGTGTGTCCGCGCTGCTGCCCTCTGAGCCCTTAACACTG 76
 61 GAAGAGCTTCATTTCTTCTTATTTTAAATGACCTTGAAGCTTCTTGAAGAT 120
 77 GAAGAGCTTCATTTCTTCTTATTTTAAATGACCTTGAAGCTTCTTGAAGAT 136
 121 CAAAAGGGCTGCGGATCTTATCAAGTTGGCCAGATGTCGATGGGCGGCAAT 180
 137 CAAAAGGGCTGCGGATCTTATCAAGTTGGCCAGATGTCGATGGGCGGCAAT 196
 181 GCGTTGGGCTTCTCACTGAGATTTCCGAGACACAGCTGAGACAGTGTGGCTTCAAC 240
 197 GCGTTGGGCTTCTCACTGAGATTTCCGAGACACAGCTGAGACAGTGTGGCTTCAAC 256
 241 CTTTCAATGCTGGGCTTGGTGGAGTGGGCAATCTGTGTGACCGGCTTCTGAGCCAG 300
 257 CTTTCAATGCTGGGCTTGGTGGAGTGGGCAATCTGTGTGACCGGCTTCTGAGCCAG 316
 301 TTCCCTTCTGGAAGGTGTATCACACTGTTCAATGATTTGGCTGGCCACATGAGTGT 360
 317 TTCCCTTCTGGAAGGTGTATCACACTGTTCAATGATTTGGCTGGCCACATGAGTGT 376
 361 TTGTCGGTGTGATCTCAAGTGAATGCTGTTTGGGGAAGTCAACTTGGCGAGTTGGTG 420
 377 TTGTCGGTGTGATCTCAAGTGAATGCTGTTTGGGGAAGTCAACTTGGCGAGTTGGTG 436
 421 GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCAACTGAGATGAGTCACTATATATC 480
 437 GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCAACTGAGATGAGTCACTATATATC 496

OY		481	TTCAACAGACTCTCCACATGAACATGATGGACAATCAAGTTCGTGCAGCCTATTGTTGGG	540
Db		497	TTCAAACAGACTCACCATBAACATGATGACATCTACGTTTGTCGACCTATTGTTGGG	556
OY		541	CTGTCGTGGCGTGGTGGCTGCCAAAGCCTCTACCCGAGGGAGCGAGAGATAAAGATCAG	600
Db		557	CTGTCGTGGCGTGGTGGCTGCCAAAGCCTCTACCCGAGGGAGCGAGAGATAAAGATCAG	616
OY		601	ACAGCAACGATACCACAGTTTGTCTGGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	660
Db		617	ACAGCAACGATACCACAGTTTGTCTGGCCATGCTGGGCGCCCTCTTCTTGTGGATGTTCTGG	676
OY		661	CCAAGTTTCAACTCGCTCTGCTGTAGAAGTCCATTCGAAAGAAAGAAAGAACCCGTCCTAAC	720
Db		677	CCAAGTTTCAACTCGCTCTGCTGTAGAAGTCCATTCGAAAGAAAGAAAGAACCCGTCCTAAC	736
OY		721	ACCTACTATGCTGTAGACAGTCAAGCGTGGTGAACAGCCATCTCAGGGTCACTCCTTGGCTCAC	780
Db		737	ACCTACTATGCTGTAGACAGTCAAGCGTGGTGAACAGCCATCTCAGGGTCACTCCTTGGCTCAC	796
OY		781	CCCCAAGGGAAGATCAGCAGACTTATGTGCACAGTGGCGGTGTGGCAGAGGCGTGGCT	840
Db		797	CCCCAAGGGAAGATCAGCAGACTTATGTGCACAGTGGCGGTGTGGCAGAGGCGTGGCT	856
OY		841	GTTGGGTACCTGGTGTACACCTGATCCCTTCTCCGATGGCTTGGCCATGGTCCGCTGGTGTG	900
Db		857	GTTGGGTACCTGGTGTACACCTGATCCCTTCTCCGATGGCTTGGCCATGGTCCGCTGGTGTG	916
OY		901	GCTGGGCTGATCTCCGTGGGGGAGACCAAGTACCTGCCGGGGGTGTGTGAACGAGTGTG	960
Db		917	GCTGGGCTGATCTCCGTGGGGGAGACCAAGTACCTGCCGGGGGTGTGTGAACGAGTGTG	976
OY		961	GGGATTCGCCACAGCTCCATCATGGGCTACAACTTCAGCTTGTGGGTCTGCTTTGAGAG	1020
Db		977	GGGATTCGCCACAGCTCCATCATGGGCTACAACTTCAGCTTGTGGGTCTGCTTTGAGAG	1036
OY		1021	ATCATCTACATTTGTGTCTGTCTGTCTGTATACCGTGGAGCCGGCAATGCAATGATTGGC	1080
Db		1037	ATCATCTACATTTGTGTCTGTCTGTCTGTATACCGTGGAGCCGGCAATGCAATGATTGGC	1096
OY		1081	TTCGAGGCTCCTCCTCAGCATTTGGGGAATCAGCTTGGCATCGATATGCTCTCAGCGCT	1140
Db		1097	TTCGAGGCTCCTCCTCAGCATTTGGGGAATCAGCTTGGCATCGATATGCTCTCAGCGCT	1156
OY		1141	GGTCTCTCTGACAGGTTTTGCTCCTTAATCTTAAAAATATGAAAAGCACTCATAGAGCTAAA	1200
Db		1157	GGTCTCTCTGACAGGTTTTGCTCCTTAATCTTAAAAATATGAAAAGCACTCATAGAGCTAAA	1216
OY		1201	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGTGATTTTAA	1254
Db		1217	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCTCATTTGGCTGTGTGATTTTAA	1270
RESULT 7				
ABO18966				
LOCUS				
DEFINITION			Homo sapiens RhDVA (FK) mRNA for Rh blood group D antigen (Rhd),	PRI 03-JUL-1999
ACCESSION			complete cds.	
VERSION			ABO18966.1 GI:5360242	
KEYWORDS			RhDVA (FK); Rh blood group D antigen (Rhd).	
SOURCE			Homo sapiens (human)	
ORGANISM			Homo sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE			1 (bases)	
JOURNAL			Hyodo,H., Ishikawa,Y., Kaeshiwee,K., Ogawa,A., Watanabe,Y.,	
REFERENCE			Tsuneyama,H., Toyoda,C., Uchikawa,M., Akaza,T. and Fujii,T.	
AUTHORS			Polymorphisms of RhDva in Japanese	
TITLE			Unpublished	
			2 (bases 1 to 1254)	
			Uchikawa,M., Hyodo,H. and Ishikawa,Y.	
			Direct Submission	

JOURNAL									
Submitted (20-OCT-1998)/Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research 4-1-3, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)									
FEATURES									
Source									
1. .1254 /organism="Homo sapiens" /mol_type="mRNA" /isolate="FK" /db_xref="taxon:9606" 1. .1254 /gene="RhDvA (FK)" 1. .1254 /gene="RhDvA (FK)" /feature="RhdvA (FK)" /note="Rhd variant RhDvA (FK)" /codon_start=1 /product="Rh blood group D antigen (Rhd)" /protein_id="PAA81899.1" /db_xref="gi:5360243" /translation="MSKRYPRSVRCPLNALTLBALILFLPFTTHDASLEDDKGI VASVQVGDVLTMAAIGLGLPTSSFRHRSSVAFFNLPMLALGVOMALLDGLSDF SGKRVITLFSRLATMSALSVLISVDVLAGVNAQLVMVLVEYTLAAGNLRVTSIN FNDYHMMMMHIIYVFAVFGSLVAMCLPKPLEGEDBDQDRAIPLSLAMGALFLM PMFVNSALLRSPIDRKXNAVENTYAAVAVSVTAISGSLHAPQKISKTYVHSAVL GVAVGTSCHLIPSPMLAMVLGVAGLISVGARVLPCCNRRVIGIPHSSIMGVNFSS LGILGEIITLYVLVDLVTVAGNMGIRPVLLSIGRLSLATLYALTLTGILLNPK WKADHEAKYFPDDQVFMKPHLAVGF" 									
ORIGIN									
Query Match 99.7%; Score 1250.8; DB 9; Length 1254;									
Matches 1252; Conservatvity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0									
QY	1	ATGAGCTTAAAGTACCCGCGGTCTGTCCGCGCGCTGCTGCCCTTCGTGGGCGCTTAACATG	60						
Db	1	ATGAGCTTAAAGTACCCGCGGTCTGTCCGCGCGCTGCTGCCCTTCGTGGGCGCTTAACATG	60						
QY	61	GAAGCAGCTCATATTCCTCTCTTATTTTAAACCACTATGACGCTTCCTTAAGAT	120						
Db	61	GAAGCAGCTCATATTCCTCTCTTATTTTAAACCACTATGACGCTTCCTTAAGAT	120						
QY	121	CAAAAGGGGCGTGGGCACTCTATCAAGATTGGCCAAAGATGACCGGTATGGCGGCATT	180						
Db	121	CAAAAGGGGCGTGGGCACTCTATCAAGATTGGCCAAAGATGACCGGTATGGCGGCATT	180						
QY	181	GGCTTGGGCTTCCTCACTCGAGTTTCCGAGACACACAGCTGAGCAGTGTGGCTTCAAC	240						
Db	181	GGCTTGGGCTTCCTCACTCGAGTTTCCGAGACACACAGCTGAGCAGTGTGGCTTCAAC	240						
QY	241	CTCTTCATGCTGGCGCTTGATGTGCAGTGGGCAATCTGTGAGCGGCTTCCTGAACCG	300						
Db	241	CTCTTCATGCTGGCGCTTGATGTGCAGTGGGCAATCTGTGAGCGGCTTCCTGAACCG	300						
QY	301	TTCCCTCTTGGGAAGGTGTGATACACATCTGTTCAATTTGGGCTGGCCACATGATGCT	360						
Db	301	TTCCCTCTTGGGAAGGTGTGATACACATCTGTTCAATTTGGGCTGGCCACATGATGCT	360						
QY	361	TTGTGCGGAGCTGATCTCACTGAGTGTCTGTGGGGAAGGTCAACTTGAGCGCAGTTGGTG	420						
Db	361	TTGTGCGGAGCTGATCTCACTGAGTGTCTGTGGGGAAGGTCAACTTGAGCGCAGTTGGTG	420						
QY	421	GTGATGCTGTGTGAGGTGACAGCTTTAAGCAACTGAGGATGTGATCAGTAATATC	480						
Db	421	GTGATGCTGTGTGAGGTGACAGCTTTAAGCAACTGAGGATGTGATCAGTAATATC	480						
QY	481	TTCAACACACACTACACATGAACTATGACATCTACGTTGTGGACGCTTATTTGGG	540						
Db	481	TTCAACACACACTACACATGAACTATGACATCTACGTTGTGGACGCTTATTTGGG	540						
QY	541	CTGTCTGTGGCTGTGGTGTGCGCAAGGCTGTAACCGGAGGAAAGGAGATTAAGATCAG	600						
Db	541	CTGTCTGTGGCTGTGGTGTGCGCAAGGCTGTAACCGGAGGAAAGGAGATTAAGATCAG	600						

QY 601 ACAGCAAGATACCAAGTTTGTCTGCATAGCTGGGCGCCCTCTTCTTGATGATTCTGG 660
DB 601 ACAGCAAGATACCAAGTTTGTCTGCATAGCTGGGCGCCCTCTTCTTGATGATTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTGCTGAGAAAGTCCATGAAAGAAAGATCCCTGTTCAAC 720
DB 661 CCAAGTTTCAACTCTGCTGCTGAGAAAGTCCATGAAAGAAAGATCCCTGTTCAAC 720
QY 721 ACCTATAGCTGTAGCAAGTCAAGCTGTGACAGACCATCTCAAGGATCCTTGGCTAC 780
DB 721 ACCTATAGCTGTAGCAAGTCAAGCTGTGACAGACCATCTCAAGGATCCTTGGCTAC 780
QY 781 CCCCAAGGAAAGATTCAGCAAGATTATGTCACAGTGCCTGTGGCAGAGCCGTGCT 840
DB 781 CCCCAAGGAAAGATTCAGCAAGATTATGTCACAGTGCCTGTGGCAGAGCCGTGCT 840
QY 841 GTGGGTAAGCTGCTGATCACTGATCCCTTCTCGGTGGCTTGCATAGTCTGGTCTTGTG 900
DB 841 GTGGGTAAGCTGCTGATCACTGATCCCTTCTCGGTGGCTTGCATAGTCTGGTCTTGTG 900
QY 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTTGTACCGAGTGTG 960
DB 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACCTGCGGGGTGTTGTACCGAGTGTG 960
QY 961 GGGATTCCCAAGCTCCATTCATGAGGCTTCAACATTCACTTGTGGTCTGTTGAGAG 1020
DB 961 GGGATTCCCAAGCTCCATTCATGAGGCTTCAACATTCACTTGTGGTCTGTTGAGAG 1020
QY 1021 ATCATCTAATTTGCTGCTGCTGCTGATATACCGTCGAGCGCGCAATGGCATATTTGGC 1080
DB 1021 ATCATCTAATTTGCTGCTGCTGCTGATATACCGTCGAGCGCGCAATGGCATATTTGGC 1080
QY 1081 TTCAGGCTCCTCCAGCATTTGGGGAATTCAGCTTGGCCATGCTGATAGCTTCAAGCTCT 1140
DB 1081 TTCAGGCTCCTCCAGCATTTGGGGAATTCAGCTTGGCCATGCTGATAGCTTCAAGCTCT 1140
QY 1141 GGTCTCTGAGAGTTTCTCTTAATTTTAAATATGAAAGAAAGCACTTATGAGGCTTAA 1200
DB 1141 GGTCTCTGAGAGTTTCTCTTAATTTTAAATATGAAAGAAAGCACTTATGAGGCTTAA 1200
QY 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTGATTTGGCTTTGATTTTAA 1254
DB 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTGATTTGGCTTTGATTTTAA 1254

RESULT 8
AB018967
LOCUS
DEFINITION Homo sapiens RhDva(TO) mRNA for Rh blood group D antigen (RHD),
complete cds.
ACCESSION AB018967
VERSION AB018967.1 GI:5360244
KEYWORDS RhDva(TO); Rh blood group D antigen (RHD).
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,
Tanueyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
TITLE Polymorphisms of RhDva in Japanese
JOURNAL Unpublished
REFERENCE
AUTHORS Uchikawa, M., Hyodo, H. and Ishikawa, Y.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1998) Hi-Jonohu Hyodo, Japanese Red Cross Central
Blood Center, Research, 4-1-31, Hi-roo, Shibuya, Tokyo 150-0012,
Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009,
Fax:+81-3-3406-7892)
FEATURES
source
1..1254
/organism="Homo sapiens"
/mol_type="mRNA"

gene
CDS
/isolate="TO"
/db_xref="taxon:9606"
1..1254
/gene="RhDva(TO)"
1..1254
/gene="RhDva(TO)"
/note="RhD variant RhDva(TO)"
/codon_start=1
/product="Rh blood group D antigen (RHD)"
/protein_id="BA01900.1"
/db_xref="GI:5360245"
/translation="MSSKYPKPRVRLCTPIALTLBAALLPYFTTHDASLEDKGL
VASVOGDLPTWMAAIGLPLPSRRHSVVARFLVLAIVGMALLDGLSOP
SGRVITLPSIRLATMSALSVLISVDVAGKNLQVVMVLVEVTLGNLMTVSNL
FNDYHNNMMHIVFPAAYGLSVANCLPRLPEGEDKQDQATIPSLAMGALPLM
FMPSFNALRSPIQRKNAMFNTYVAIVAVVTALISGLSARPOKISKITYHSAVLA
GGVAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLPCCNRRVIGIIPHSISMGVRS
LGLLGEIYIVLVLDIVGAGMIGFVULSISGLSLAIVALTSLTLGLMLTKI
MKAPHEAKTFDDQVFWKPHLAVGF"

ORIGIN

Query Match 99.7%; Score 1250.8; DB 9; Length 1254;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGGGCTGTGTCCGGCGCTGCTGCCCTTGGGCCCTTAACACTG 60
DB 1 ATGAGCTTAAGTACCCGGGCTGTGTCCGGCGCTGCTGCCCTTGGGCCCTTAACACTG 60
QY 61 GAAGCAGCTCTGATCTCCTCTTCTATTTTAAATCCCACTAAGAGCTTCTTGAAGAGAT 120
DB 61 GAAGCAGCTCTGATCTCCTCTTCTATTTTAAATCCCACTAAGAGCTTCTTGAAGAGAT 120
QY 121 CAAAAGGGCTGTGGCATCTTATCAAGTTGGCCAAAGATCTGACCGTGATGGCCCATY 180
DB 121 CAAAAGGGCTGTGGCATCTTATCAAGTTGGCCAAAGATCTGACCGTGATGGCCCATY 180
QY 181 GGGTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
DB 181 GGGTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
QY 241 CTCTTCATGCTGGCGCTTGTGTGCAAGGCAATCTGCTGCAAGGCTTCTTGAAGCAG 300
DB 241 CTCTTCATGCTGGCGCTTGTGTGCAAGGCAATCTGCTGCAAGGCTTCTTGAAGCAG 300
QY 301 TTCCCTTGGGAAAGGTGATCATCACTGTTCAATTTGGCTGGCCACCATGAGTGTCT 360
DB 301 TTCCCTTGGGAAAGGTGATCATCACTGTTCAATTTGGCTGGCCACCATGAGTGTCT 360
QY 361 TTGTGGTGTGATCTGATGATGATGCTGTCTTGGGGAAGGTCAACTTGGGCGAGTTGGTG 420
DB 361 TTGTGGTGTGATCTGATGATGATGCTGTCTTGGGGAAGGTCAACTTGGGCGAGTTGGTG 420
QY 421 GTGATGGCTGTGGTGGAGGTGACAGCTTTTGGCAACTTGAAGATGTCATCAATATATC 480
DB 421 GTGATGGCTGTGGTGGAGGTGACAGCTTTTGGCAACTTGAAGATGTCATCAATATATC 480
QY 481 TTCAACACAGCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTCAACACAGCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGTGGCTGTGGCTGTGGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG 600
DB 541 CTGTCTGTGGCTGTGGCTGTGGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG 600
QY 601 ACAGCAAGATACCAAGTTTGTCTGCATAGCTGGGCGCCCTCTTCTTGATGATTCTGG 660
DB 601 ACAGCAAGATACCAAGTTTGTCTGCATAGCTGGGCGCCCTCTTCTTGATGATTCTGG 660
QY 661 CCAAGTTTCAACTCTGCTGCTGAGAAAGTCCATGAAAGAAAGATCCCTGTTCAAC 720
DB 661 CCAAGTTTCAACTCTGCTGCTGAGAAAGTCCATGAAAGAAAGATCCCTGTTCAAC 720

Qy		721	ACCACTAATGCTAGAGCATGAGGTGGTGA CAGCATTCTAGGGTCATCTTGAGCTCAC	780
Db		721	ACCTACTATGCCTGTAAGCATGAGGTGGTGA CAGCATTCTAGGGTCATCTTGAGCTCAC	780
Qy		781	CCCCAAGGAAAGATCAGCAAGACTTATGTGACA GAGTGCAGTGTGGCAGAGAGCGTGCT	840
Db		781	CCCCAAGGAAAGATCAGCAAGACTTATGTGACA GAGTGCAGTGTGGCAGAGAGCGTGCT	840
Qy		841	GTTGGTACCTTGCTGTCACTGTATCCCTTCTCC GTGAGCTTGGCAGATGTCGTGGTCTTG	900
Db		841	GTTGGTACCTTGCTGTCACTGTATCCCTTCTCC GTGAGCTTGGCAGATGTCGTGGTCTTG	900
Qy		901	GCTGGGCGTGAATCTCCGTGGGGGGAGCCAA GATACCTGCCGGGGGTGTGTAAACCGAGTGTG	960
Db		901	GCTGGGCGTGAATCTCCGTGGGGGGAGCCAA GATACCTGCCGGGGGTGTGTAAACCGAGTGTG	960
Qy		961	GGGATATCCCACAGAGCTCCATCATGAGGACTA CAAGCTTCAGCTTGCTGTGGCTTGGAGAG	1020
Db		961	GGGATATCCCACAGAGCTCCATCATGAGGACTA CAAGCTTCAGCTTGCTGTGGCTTGGAGAG	1020
Qy		1021	ATCATCTACATTTGCTGCTCTGTGCTTGATA CCCTGGAGCCGGCAATGSCATGATTTGGC	1080
Db		1021	ATCATCTACATTTGCTGCTCTGTGCTTGATA CCCTGGAGCCGGCAATGSCATGATTTGGC	1080
Qy		1081	TTCACAGGCTCCCTCAGCATTTGGGGAATCGA CTGGCATTCGTGATAGCTCTCAAGCTCT	1140
Db		1081	TTCACAGGCTCCCTCAGCATTTGGGGAATCGA CTGGCATTCGTGATAGCTCTCAAGCTCT	1140
Qy		1141	GGTCCTCTGACAGAGTTTGCTCCTTAATCTTAA ATATGAGAAGCACTCATGAGGCTTAA	1200
Db		1141	GGTCCTCTGACAGAGTTTGCTCCTTAATCTTAA ATATGAGAAGCACTCATGAGGCTTAA	1200
Qy		1201	TATTTTGATGACCAAAGTTTTCTGGAAGTTTCT CATTTGSGCTGTGGATTTTAA	1254
Db		1201	TATTTTGATGACCAAAGTTTTCTGGAAGTTTCT CATTTGSGCTGTGGATTTTAA	1254
RESULT 9				
AF510069			1480 bp mRNA linear PRI 27-NOV-2002	
LOCUS			Homo sapiens Rhesus blood group D antigen D(667) variant mRNA,	
DEFINITION			complete cds.	
ACCESSION			AF510069	
VERSION			AF510069.1 GI:25573107	
KEYWORDS			.	
SOURCE			Homo sapiens (human)	
ORGANISM			Homo sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE			1 (bases 1 to 1480)	
JOURNAL			Notiz.-Piemme,F., Lee,K., Le Pennec,P.Y., Simon,P., Kazup,P.,	
PUBMED			Bachir,D., Rouzaud,A.M., Rousseil,M., Juszczak,G., Menanteau,C.,	
AUTHORS			Rouger,P., Koch,R., Cartron,J.P. and Ansart-Piemme,H.	
TITLE			Rare RHCE phenotypes in black individuals of Afro-Caribbean origin:	
JOURNAL			Identification and transfusion safety	
PUBMED			Blood 100 (12), 4227-4231 (2002)	
AUTHORS			2 (bases 1 to 1480)	
TITLE			Notiz.-Piemme,F., Cartron,J.-P. and Ansart-Piemme,H.	
JOURNAL			Direct Submission	
PUBMED			Submitted (08-MAY-2002) CNRGS, INTS, 20 rue Bouvier, Paris 75011,	
AUTHORS			France	
FEATURES			Location/Qualifiers	
source			1..1480	
CDS			/organism="Homo sapiens"	
			/mol_type="mRNA"	
			/db_xref="taxon:9606"	
			/note="Isolated from Black individual of Afro-Caribbean	
			origin"	
			44..1297	
			/note="RHD"	
			/codon_start=1	
			/product="Rhesus blood group D antigen D(667) variant"	

Query Match	Similarity	Score	DB	Length
Best Local Match	99.8%	1250.8	9	1480
Match	Conservative	0	Mismatches	2
			Indels	0
			Gaps	0

Query	Match	Similarity	Score	DB	Length
1	ATGAGCTCTAAGTACCCCGGCTCTGTCGGGCGCTCTGCCCTCTGGGCGCTTAACACTG	60			
44	ATGAGCTCTAAGTACCCCGGCTCTGTCGGGCGCTCTGCCCTCTGGGCGCTTAACACTG	103			
61	GAAGCAGCTCTCATCTCTCTCTTCTATTATTTTACCACTATAGCGCTTCTTTAAGAGAT	120			
104	GAAGCAGCTCTCATCTCTCTCTTCTATTATTTTACCACTATAGCGCTTCTTTAAGAGAT	163			
121	CAAAAGGGGCTCTGGGCAATCCATTAATGGCCAAAGATCTGAACCGTGAATGGCGGCATT	180			
164	CAAAAGGGGCTCTGGGCAATCCATTAATGGCCAAAGATCTGAACCGTGAATGGCGGCATT	223			
181	GGCTTGGGCTCTCTCACTCGAGTTTCCGAGACACAGCTGGAGAGATGTGGCTTCAAC	240			
224	GGCTTGGGCTCTCTCACTCGAGTTTCCGAGACACAGCTGGAGAGATGTGGCTTCAAC	283			
241	CTTTGATGCTGGCGCTTGTGTGTGAGTGGGCAATCTTGTGAGCGGCTTCTTGAGCCAG	300			
284	CTTTGATGCTGGCGCTTGTGTGTGAGTGGGCAATCTTGTGAGCGGCTTCTTGAGCCAG	343			
301	TTCCCTTTGGGGAAGTGTGATCACTATGTTCAATTTCCGCTGGCCACCATATGATGCT	360			
344	TTCCCTTTGGGGAAGTGTGATCACTATGTTCAATTTCCGCTGGCCACCATATGATGCT	403			
361	TTGTGCGTGTGATCTCACTGAGTGTCTGCTTGGGGAAGTGTCACTTGGCGAGTGTGTG	420			
404	TTGTGCGTGTGATCTCACTGAGTGTCTGCTTGGGGAAGTGTCACTTGGCGAGTGTGTG	463			
421	GTGATGTGTCTGTGTGAGTGTGACAGCTTAAAGGCACTGAGAGTGTATCACTAATATC	480			
464	GTGATGTGTCTGTGTGAGTGTGACAGCTTAAAGGCACTGAGAGTGTATCACTAATATC	523			
481	TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	540			
524	TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	583			
541	CTGTCTGTGGCTGTGTGTGCTGCAAAAGCTTCAACCGGGAAGGGAAGGGAATTAAGTCA	600			
584	CTGTCTGTGGCTGTGTGTGCTGCAAAAGCTTCAACCGGGAAGGGAAGGGAATTAAGTCA	643			
601	ACAGCAAGATACCAAGTTTGTCTGTGCAATGCTGTGGCGGCTTCTTGTGATGTTCTG	660			
644	ACAGCAAGATACCAAGTTTGTCTGTGCAATGCTGTGGCGGCTTCTTGTGATGTTCTG	703			
661	CCAAAGTTCAACTGTCTGTGTGCAAAAGTCCATTCGAAGGAAGATATCCGTTTCAAC	720			
704	CCAAAGTTCAACTGTCTGTGTGCAAAAGTCCATTCGAAGGAAGATATCCGTTTCAAC	763			
721	ACCTACTATGCTGTGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG	780			
764	ACCTACTATGCTGTGAGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG	823			
781	CCCAAGGAAGATCAGCAAGATTTATGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT	840			
824	CCCAAGGAAGATCAGCAAGATTTATGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT	883			
841	GTGGTACTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900			

Db 1004 GGGATTCCTCCAGCTTCATCATGGGCTACCACTTCAGCTTGGCTGCTGGAGAG 1063
Qy 1021 ATCATCTACATTTGCTGCTGCTGCTGATATACCGTGGAGCCGCAATGAGCATGATGGC 1080
Db 1064 ATCATCTACATTTGCTGCTGCTGCTGATATACCGTGGAGCCGCAATGAGCATGATGGC 1123
Qy 1081 TTCAGGCTCTCTCTCAGCATTTGGGGAATCAGCTTGGCCATCGTATAGCTTCACGCTCT 1140
Db 1124 TTCAGGCTCTCTCTCAGCATTTGGGGAATCAGCTTGGCCATCGTATAGCTTCACGCTCT 1183
Qy 1141 GGTCTCTCTGACAGTTTGTCTCTTAATCTTAAATATAGAAAGCACTCTATAGGCTTAA 1200
Db 1184 GGTCTCTCTGACAGTTTGTCTCTTAATCTTAAATATAGAAAGCACTCTATAGGCTTAA 1243
Qy 1201 TATTTGATGACCAAGTTTCTGAGATTCTCATTTGGCTGTTGATTTTAA 1254
Db 1244 TATTTGATGACCAAGTTTCTGAGATTCTCATTTGGCTGTTGATTTTAA 1297

RESULT 11
AJ784309 1251 bp DNA linear PRI 25-JUL-2004
LOCUS AJ784309
DEFINITION Homo sapiens partial RHD gene for rhesus blood group D antigen,
weak D allele, exon 4.
ACCESSION AJ784309
VERSION AJ784309.1 GI:50657032
KEYWORDS RHD gene; rhesus blood group D antigen; weak D allele.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Doescher A., Wagner F., Schunter F. and Petershofen E.K.
Four new RHD alleles with previously unknown polymorphism
Unpublished
2 (bases 1 to 1251)
Doescher A.
Direct Submission
Submitted (19-JUL-2004) Doescher A., Molecular Diagnostics, German
Red Cross Blood Transfusion Centre, Brandenburger Str. 21, D-26133
Oldenburg, GERMANY
Related sequence: L08429.
Location/Qualifiers
1. 1251
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="141/2003"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.2-p34"
/cell_type="erythrocyte"
/tissue_type="peripheral blood"
1. 1251
/gene="RHD"
/gene="RHD"
/gene="RHD"
/codon_start=1
/allele="weak D"
/product="rhesus blood group D antigen"
/protein_id="CAH04919.1"
/db_xref="GI:50657032"
/translation="MSSKPSRVRCCLPLMALTEALILLPFFTHYDLSLDOKGL
VASVOGDLTMAAIGLITSPRHSMSVAFLPMLALGVMAILLDGLSOP
SGRVVITLPSIRLATMSALSVLSVDAVLCKNALDQLVVVVLEVRALGNLVISNI
FNDYHNKMKQIIVFAVFGLSVAMCLPKPLPGSTBKDQATIPSLISALGALFLMM
FMSFNSALRSPIERKNAFNTYVAVAVSIVVLAIGSSLAHQGKISKYVNSAVLA
GVAIVGSHLIPSPMLAVLGAGLISVGAKYLPGCCNRYLGIPIHSSIMGYNSL
LGILGELIYLVLDVTVAGNGMIGFQVLISIGELSLAIVIALTSGLLTLNLIKI
WKAHBAKXFDDQVFWKPHLAVGF"
exon 1. 1251
/gene="RHD"
/number=4
variation 513

ORIGIN
Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/gene="RHD"
/note="single nucleotide polymorphism, SNP"
/replace="c"
Qy 1 ATGAGCTTAAGTACCGGGGCTGTCCGGCGGTGCTGCCCTTGGGCGCTTAACATG 60
Db 1 ATGAGCTTAAGTACCGGGGCTGTCCGGCGGTGCTGCCCTTGGGCGCTTAACATG 60
Qy 61 GAAGCACTCTCATCTCTCTCTCTATATTTTATACCACATAGACGCTTCTTAGAGAT 120
Db 61 GAAGCACTCTCATCTCTCTCTCTCTATTTTATACCACATAGACGCTTCTTAGAGAT 120
Qy 121 CAAAAGGGGCTCGTGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGGCGCAT 180
Db 121 CAAAAGGGGCTCGTGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGGCGCAT 180
Qy 181 GGCTTGGGCTCTCTCACTCGAGTTTCCGGAGACAGCTGGAGAGTGGCTTCAAC 240
Db 181 GGCTTGGGCTCTCTCACTCGAGTTTCCGGAGACAGCTGGAGAGTGGCTTCAAC 240
Qy 241 CTCTTATGCTGGCGCTTGGTGTGCAAGTGGCAATCTGCTGAGCGCTTCTGAGCCAG 300
Db 241 CTCTTATGCTGGCGCTTGGTGTGCAAGTGGCAATCTGCTGAGCGCTTCTGAGCCAG 300
Qy 301 TTCTTATGCTGGAGGTGTGTCATCACTGTTCAATTTGGCTGGCCACATAGTCT 360
Db 301 TTCTTATGCTGGAGGTGTGTCATCACTGTTCAATTTGGCTGGCCACATAGTCT 360
Qy 361 TTGTGCGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCAACTTGGCGAGTTGGT 420
Db 361 TTGTGCGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCAACTTGGCGAGTTGGT 420
Qy 421 GTGATGCTGTGTGAGGTGACAGCTTTAGCAACTGAGATGTGATCAATATATC 480
Db 421 GTGATGCTGTGTGAGGTGACAGCTTTAGCAACTGAGATGTGATCAATATATC 480
Qy 481 TTCAACACAGACTACCAATGAACATGATGACATCTTATGCGAGCTTATTTTGGG 540
Db 481 TTCAACACAGACTACCAATGAACATGATGACATCTTATGCGAGCTTATTTTGGG 540
Qy 541 CTGTCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CTGTCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 601 ACAGCAACGATACCAAGTTTGTCTGCAATGCTGGGCGCTCTTCTTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCAAGTTTGTCTGCAATGCTGGGCGCTCTTCTTGTGATGTTCTGG 660
Qy 661 CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAATGGAAGAAATGCGTGTTCAC 720
Db 661 CCAAGTTTCAACTCTGCTCTGCTGAGAGTCCAATGGAAGAAATGCGTGTTCAC 720
Qy 721 ACTTACTATGCTGTACAGTCAAGCGTGTGACAGCCATCTCAAGGTCAATCTTGGCTCAC 780
Db 721 ACTTACTATGCTGTACAGTCAAGCGTGTGACAGCCATCTCAAGGTCAATCTTGGCTCAC 780
Qy 781 CCCCAAGGGAAGTACGAAAGTTATATGTCACAGTGGCGGTGTTGGCAGAGCGTGGCT 840
Db 781 CCCCAAGGGAAGTACGAAAGTTATATGTCACAGTGGCGGTGTTGGCAGAGCGTGGCT 840
Qy 841 GTGGGTAACCTCGTGTCACTGATCCCTTCTCGGTGCTTGCATGCTGCTGCTTGTG 900
Db 841 GTGGGTAACCTCGTGTCACTGATCCCTTCTCGGTGCTTGCATGCTGCTGCTTGTG 900
Qy 901 GCTGGGCTATCTCGCTCGGGGAGGCAAGTACTGCGCGGGGTGTTACCGAAGTCTG 960
Db 901 GCTGGGCTATCTCGCTCGGGGAGGCAAGTACTGCGCGGGGTGTTACCGAAGTCTG 960

QY	961	GGGATTTCCCAAGCTCCATCATGAG3CTPACAACTTACGTTGCTGGGCTCTGTTGAGAG	1020
Db	961	GGGATTTCCCAAGCTCCATCATGAG3CTPACAACTTACGTTGCTGGGCTCTGTTGAGAG	1020
QY	1021	ATCATCTACATGTGCTGCTGAGTCTTGATACCGTGGAGGCCGGAATGGGATGATGGC	1080
Db	1021	ATCATCTACATGTGCTGCTGAGTCTTGATACCGTGGAGGCCGGAATGGGATGATGGC	1080
QY	1081	TTCCAGGTCCTTCCTAGCATTTGGGGAACCTCAGCTTGGCCATGTGATAGCTCTCAGCTCT	1140
Db	1081	TTCCAGGTCCTTCCTAGCATTTGGGGAACCTCAGCTTGGCCATGTGATAGCTCTCAGCTCT	1140
QY	1141	GCTCCTCAGAGGTTTGCTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA	1200
Db	1141	GCTCCTCAGAGGTTTGCTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA	1200
QY	1201	TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTT	1251
Db	1201	TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTT	1251

RESULT 12

LOCUS AJ784310

DEFINITION Homo sapiens partial RHD gene for rhesus blood group D antigen, weak D allele, exon 9.

ACCESSION AJ784310

VERSION AJ784310.1 GI:50657034

KEYWORDS RHD gene; rhesus blood group D antigen; weak D allele.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Doeschner A., Wagner F., Schunter F. and Peterahofen E.K.

TITLE Four new RHD alleles with previously unknown polymorphism

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1251)

AUTHORS Doeschner A.

TITLE Direct Submission

JOURNAL Submitted (19-JUL-2004) Doeschner A., Molecular Diagnostics, German Red Cross Blood Transfusion Centre, Brandenburger Str. 21, D-26133 Oldenburg, GERMANY

COMMENT Related sequence: L08429.

FEATURES

source Location/Qualifiers

1..1251

```

gene
/ncbi_type="genomic DNA"
/isolate="155/2003"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.2-p34"
/cell_type="erythrocyte"
/tissue_type="peripheral blood"
1..1251
/gene="RHD"
<1..>1251
/gene="RHD"
/codon_start=1
/allele="weak D"
/product="rhesus blood group D antigen"
/protein_id="CAH04920.1"
/db_xref="GI:50657035"
/translation="MSSKYPSPSRVRCCLPMLALTEALILFYPPHYDASLEDOQKGL
VASVQGOGLITMAAIGAGFLTSSFRHSMSSVAENLPMLAGVQMAILLDGLSOPP
SKRVITLPSILATMSALSVLISVDALGVKNLAQLVNLVEVTLAIGNIRVTSINI
FNTDTHMMNNLIVYPAAYFGSLVAMCPLPEEGEDKQDQATISLSLMAAGFLMMW
PSPNSALMRSLPIERKNAVENTYYAAVAVSVTLAISGSLAHPOGKISKTYMGALVLA
GVAVGTSCILPSLPMALVGLVAGLISVGAKYLPCCCNRVIGITHSSIMGNPSL
IGLLEIIVTILVLDYAGAGNGMTGFOVLLSIGELSLATVIALTSGLTGLLNLNKI
WKAHPKATFDDQVIMKTPHLAVGF"
1..1251
/gene="RHD"

```

```
variation
/number=9
1221
/gene="RHD"
/note="single nucleotide polymorphism, SNP"
/replace="C"

ORIGIN
```

Query Match	99.5%	Score 1247.8;	DB 9;	Length 1251;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1249;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0;

QY	1	ATAGAGCTAAAGTACCCGCGGCTCTCTCCGCGGCTGCTCCCTCTCTGGGCGCTTAACATCG	60
Dp	1	ATAGAGCTTAAGTACCCGCGGCTCTCTCCGCGGCTGCTCCCTCTCTGGGCGCTTAACATCG	60
QY	61	GAAGCAGCTCTGATCTCTCTCTTAATTTTTTAAACCACTAATGACGCTTCTCTTAGAGAT	120
Dp	61	GAAGCAGCTCTGATCTCTCTCTTAATTTTTTAAACCACTAATGACGCTTCTCTTAGAGAT	120
QY	121	CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTCTGACCGGTAGTGGCGGCATT	180
Dp	121	CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTCTGACCGGTATGGCGGCATT	180
QY	181	GGCTTGGGCTTCTCAACCTCGAGTTTCCGGAGACAACACTGAGACAGTGTGCGCTTCAAC	240
Dp	181	GGCTTGGGCTTCTCAACCTCGAGTTTCCGGAGACAACACTGAGACAGTGTGCGCTTCAAC	240
QY	241	CTCTTCATGCTGGCGCTTGATGTGCAGTGGGCAATCCTGCTGGAACGCGCTTCTGAGCAG	300
Dp	241	CTCTTCATGCTGGCGCTTGATGTGCAGTGGGCAATCCTGCTGGAACGCGCTTCTGAGCAG	300
QY	301	TTCCCTTCTGGGAAGTGGTCAATCACATGTTCAAGTATTCGGCTGGCCACATGAGTCT	360
Dp	301	TTCCCTTCTGGGAAGTGGTCAATCACATGTTCAAGTATTCGGCTGGCCACATGAGTCT	360
QY	361	TTGTGCGTGTGATCTCAGTGGATGCTGCTTTGGAGAAAGTCAACTTGGGCGAGTTGGTG	420
Dp	361	TTGTGCGTGTGATCTCAGTGGATGCTGCTTTGGAGAAAGTCAACTTGGGCGAGTTGGTG	420
QY	421	GTGATGTGTCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGATGTCAATCAGTAATATC	480
Dp	421	GTGATGTGTCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGATGTCAATCAGTAATATC	480
QY	481	TTCAACAACAGCTACCCATGAAATATATGACACTTAAGTGTTCGAGCTTATTTGGG	540
Dp	481	TTCAACAACAGCTACCCATGAAATATATGACACTTAAGTGTTCGAGCTTATTTGGG	540
QY	541	CTGTCTGTGGCGTGGTGGCTTGCCAAACCTCTACCCGAGGGAAACGAGAGATTAAGATCAG	600
Dp	541	CTGTCTGTGGCGTGGTGGCTTGCCAAACCTCTACCCGAGGGAAACGAGAGATTAAGATCAG	600
QY	601	ACAGCAACGATACCCAGTTTGTCTGCGATGCTGGGCGCCTCTTCTTGTGATGTCTGG	660
Dp	601	ACAGCAACGATACCCAGTTTGTCTGCGCATCTGGGCGCCTCTTCTTGTGATGTCTGG	660
QY	661	CCAAGTTTCAACTGTGCTCTGTGAGAAAGTCAATGGAAGAAAGAAATGCCGTTCAC	720
Dp	661	CCAAGTTTCAACTGTGCTCTGTGAGAAAGTCAATGGAAGAAAGAAATGCCGTTCAC	720
QY	721	ACCTACTATGCTGTAGCAGTCAAGCGTGGTGAACGCACTCAGAGTCAATCTTGTGCTCAC	780
Dp	721	ACCTACTATGCTGTAGCAGTCAAGCGTGGTGAACGCACTCAGAGTCAATCTTGTGCTCAC	780
QY	781	CCCCAAGGAAAGATCAGCAAGACTTATGTGCACAGTGGCGGTGTGGCAGAGGCGTGGCT	840
Dp	781	CCCCAAGGAAAGATCAGCAAGACTTATGTGCACAGTGGCGGTGTGGCAGAGGCGTGGCT	840
QY	841	GTGGGTATCTCTGTACCTGATCCCTTCTTCGGTGGCTTGGCAATGGTGTGGGTCTTGGG	900
Dp	841	GTGGGTATCTCTGTACCTGATCCCTTCTTCGGTGGCTTGGCAATGGTGTGGGTCTTGGG	900
QY	901	GCTGGGCTGATCTCCGTGGGGAGGCAAGTACCTGCGGGGTGTGTGAACGAGTGTGG	960

Db	901	GCTGGGCTGATCTCCGTCGGGGGAGCCAAAGTACTCTGCCGGGTGTGTTAAACCGAGTGTG	960
Qy	961	GGGATTTCCCAACAAGCTCCATCATATGGGCTACAACATTCACTTGTCTGGGCTCTCTTGAGAG	1020
Db	961	GGGATTTCCCAACAAGCTCCATCATATGGGCTACAACATTCACTTGTCTGGGCTCTCTTGAGAG	1020
Qy	1021	ATCATCTACATATGTGCTGTGGTGTATATACCGTCGAGGCGGGAATGGGATGATTTGGC	1080
Db	1021	ATCATCTACATATGTGCTGTGGTGTATATACCGTCGAGGCGGGAATGGGATGATTTGGC	1080
Qy	1081	TTCCAGTCTCTCTCAGCATTGGGGAACTCAGCTTGGCCATCTGTGATAGCTCTCAAGCTCT	1140
Db	1081	TTCCAGGCTCTCTCAGCATTGGGGAACTCAGCTTGGCCATCTGTGATAGCTCTCAAGCTCT	1140
Qy	1141	GGTCCTCCGATGAGGTTTCTCCTTAAATCTTAAATAATGAAAGCACTCATGAGGCTTAA	1200
Db	1141	GGTCCTCCGATGAGGTTTCTCCTTAAATCTTAAATAATGAAAGCACTCATGAGGCTTAA	1200
Qy	1201	TATTTTGATGACCAAGATTTTCTGGAAGTTTCTCATTTGGCTGTGGATTT	1251
Db	1201	TATTTTGATGACCAAGATTTTCTGGAAGTTTCTCATTTGGCTGTGGATTT	1251

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1	Doeschner, A., Wagner, F., Schunter, F. and Peterahofen, E.K.	Four new RHD alleles with previously unknown polymorphism	Unpublished	2	(bases 1 to 1251)	Doeschner, A.	Direct Submission
1	Submitted (19-JUL-2004)	Doeschner A., Molecular Diagnostics, German Red Cross Blood Transfusion Center, Brandenburg Str. 21, D-26133 Oldenburg, GERMANY	Related sequence: L068429.	1.	1251	Location/Qualifiers	
		/organism="Homo sapiens"				/mol_type="genomic DNA"	
		/isolate="19/2004"				/db_xref="taxon:9606"	
		/chromosome="1"				/map="1p36.2-p34"	
		/cell_type="erythrocyte"				/tissue_type="peripheral blood"	
		1..1251				/gene="RHD"	
		<1..>1251				/gene="RHD"	
		/codon_start=1				/allele="weak D"	
		/product="rhesus blood group D antigen"				/protein_id="CAH04921.1"	
		/db_xref="GI:50657037"				/translation="MSKYPKRSVRCLPLMALTLLEAAILLFFPTHYDASLEDDKGLVASVVGODLTVMAIGLIGFLTSSFRRSMSVAFLFMALIGVOMAILDDGPLSQPESGKVVITLPSIRLATMSALSVLISVDNIAKGVNLAOLVYVWLVYVTLGILRMVTSNITNTIYHMMHMTIYFPAATFGISVAMCLPKPIPESTBEDDQVATTPISLHMLGALFLMMFWPFSNSALHSPLERKNAVENTYAAVASVVTALISGSLAHPOGKISKTYVSAVAIVAGVAVGWSCHLIPSPWLMVNLVAGLISVGAVYLPCCCRNVIGIPHSISMGYNFSLGLGLIGLIIYIVLIVDTVAGNCGMIGFVVLISIGLSIAIVIALITSGLTGLGLINLTKIKWAKHEAVYFDQVFMVKFPLAVGF"	

exon	1. .1251	/gene="RHD"	/number=4	520	/gene="RHD"	/note="single nucleotide polymorphism, SNP"	/replace="g"
variation							
ORIGIN							
Query Match	99.5%;	Score 1247.8;	DB 9;	Length 1251;			
Best Local Similarity	99.8%;	Pred. No. 0;					
Matches 1249;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps
QY	1	ATGAGCTTAAAGTACC	GCGGCTGTGTC	CGCGCGCTG	CCCTCTG	GGGCGCTTAA	CACGTG 60
DB	1	ATGAGCTTAAAGTACC	GCGGCTGTGTC	CGCGCGCTG	CCCTCTG	GGGCGCTTAA	CACGTG 60
QY	61	GAAGAGCTCATTTCT	CTCTCATTTTTTT	TAAACCACTA	TGAGAGCTT	CCCTTA	AGAGT 120
DB	61	GAAGAGCTCATATTCT	CTCTCATTTTTTT	TAAACCACTA	TGAGAGCTT	CCCTTA	AGAGT 120
QY	121	CAAAAGGGGCTCGT	GGCATCCTATCA	ATTGGCCAA	AGATCTGAC	CCGTATG	CGCGCAAT 180
DB	121	CAAAAGGGGCTCGT	GGCATCCTATCA	ATTGGCCAA	AGATCTGAC	CCGTATG	CGCGCAAT 180
QY	181	GCGTTGGGCTTCTCA	CTCGAGTTTCC	GAGACACAC	AGCTTGAG	CAGATGTG	CCCTTCAAC 240
DB	181	GCGTTGGGCTTCTCA	CTCGAGTTTCC	GAGAGACAC	AGCTTGAG	CAGATGTG	CCCTTCAAC 240
QY	241	CTCTTCAAGTCGGG	CTTGGTGTGCA	TGGGCAATCT	CTGTGAC	AGGCTTCT	GTAGCCAG 300
DB	241	CTCTTCAAGTCGGG	CTTGGTGTGCA	TGGGCAATCT	CTGTGTGAC	AGGCTTCT	GTAGCCAG 300
QY	301	TTCCCTTCTGGGA	AGGTGGTATCA	CACTGTTC	AGTATTC	GGCGCACA	TCAGATGCT 360
DB	301	TTCCCTTCTGGGA	AGGTGGTATCA	CACTGTTC	AGTATTC	GGCGCACA	TCAGATGCT 360
QY	361	TTTGTGCGTGTGAT	TCTCAGTGA	TGCTGTCTT	GGGGAAGTCA	CTTGGCGCA	GTGATTTGGTG 420
DB	361	TTTGTGCGTGTGAT	TCTCAGTGA	TGCTGTCTT	GGGGAAGTCA	CTTGGCGCA	GTGATTTGGTG 420
QY	421	GTGATGTGCTGTG	AGAGGTGAC	AGCTTTA	GGCAACTG	AGAGATGT	CATCACTAATATC 480
DB	421	GTGATGTGCTGTG	AGAGGTGAC	AGCTTTA	GGCAACTG	AGAGATGT	CATCACTAATATC 480
QY	481	TTTCAACACAGACT	ACACATGA	ATGACATCTA	CGTGTG	CGACCTAT	TTTGTGG 540
DB	481	TTTCAACACAGACT	ACACATGA	ATGACATCTA	CGTGTG	CGACCTAT	TTTGTGG 540
QY	541	CTGTCTGTGGCTGT	GTGCTGCCAA	AGGCTCTA	CCCGAGGGA	CCGAGATTA	AAATATCG 600
DB	541	CTGTCTGTGGCTGT	GTGCTGCCAA	AGGCTCTA	CCCGAGGGA	CCGAGATTA	AAATATCG 600
QY	601	ACAGCAAGATACCA	GTGTTGTG	CCACTG	CTGGGCGG	CCCTCTT	TTGTGATGTTCTGG 660
DB	601	ACAGCAAGATACCA	GTGTTGTG	CCACTG	CTGGGCGG	CCCTCTT	TTGTGATGTTCTGG 660
QY	661	CCAAAGTTCAA	CTGTCTGTG	CTGAGAA	GTCCCAT	CGAAGAGAA	TGCGGTGTTCAAC 720
DB	661	CCAAAGTTCAA	CTGTCTGTG	CTGAGAA	GTCCCAT	CGAAGAGAA	TGCGGTGTTCAAC 720
QY	721	ACCTACTATGTGTGA	CGATCAGG	TGTGTGACAG	CACTCAGAG	GTCAATCC	TGTGGCTAC 780
DB	721	ACCTACTATGTGTGA	CGATCAGG	TGTGTGACAG	CACTCAGAG	GTCAATCC	TGTGGCTAC 780
QY	781	CCCGAAGGGAAGAT	CAGCAAGCT	TATATGTG	CACAGTGG	GTGTTGG	CGAGGCGTGGCT 840
DB	781	CCCGAAGGGAAGAT	CAGCAAGCT	TATATGTG	CACAGTGG	GTGTTGG	CGAGGCGTGGCT 840
QY	841	GTGGGTACTGTGTG	CACTGTATCC	CTTCTCC	GTGGCTT	GCACATG	GTCTGGATTTTGTG 900
DB	841	GTGGGTACTGTGTG	CACTGTATCC	CTTCTCC	GTGGCTT	GCACATG	GTCTGGATTTTGTG 900

QY 901 GCTGGGCTGATCTCCGTCGGGGAGGCCAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGAGGCCAAGTACCTGCGGGGTGTTGTAACCGAGTCTG 960
QY 961 GGGATTCCCCAGAGCTCCATCATAGGGCTACAACTTCAAGCTTGCGGTCTGCTTGAGAG 1020
Db 961 GGGATTCCCCAGAGCTCCATCATAGGGCTACAACTTCAAGCTTGCGGTCTGCTTGAGAG 1020
QY 1021 ATCATCTACATTGTGCTGTGTGTGTAACCGTCGAGCGGGCAATGGCAATGTGGC 1080
Db 1021 ATCATCTACATTGTGCTGTGTGTGTAACCGTCGAGCGGGCAATGGCAATGTGGC 1080
QY 1081 TTCAGAGTCTCTCTCAGCATATGGGCACTCAGCTTGCCATGTGATAGCTCTCAGCTCT 1140
Db 1081 TTCAGAGTCTCTCTCAGCATATGGGCACTCAGCTTGCCATGTGATAGCTCTCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTCATAGAGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTCATAGAGCTTAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251

RESULT 14

AY449381

LOCUS AY449381 1251 bp mRNA linear PRI 03-DEC-2003
DEFINITION Homo sapiens rhesus blood group D antigen DFE mRNA, partial cde.
ACCESSION AY449381
VERSION AY449381.1 GI:38565510
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1251)
TITLE Noizat-Pirenne, F. and Ansart-Pirenne, H.
JOURNAL Direct Submission
Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
FEATURES
Source Location/Qualifiers
1..1251

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1251
/note="weak D DFE"
/codon_start=1
/product="rhesus blood group D antigen DFE"
/protein_id="AAR24078.1"
/db_xref="GI:38565511"
/translation="MSKSRPSVRCLPLMALTLLEALILLYFTHYDASLEDQGL
VASVYQODLTTHMAIGLEFLISFRHSMSVAENLPMALGVOMAILDGLFLOFP
SGKVVITLTSIRLATWSALSVALISDAVAGKVALADLVVMEVETALNENLMSVSNL
FNTDYHNNMHIYFAFAGLSVAMCLPKPEDEGDKQATVIBSLMSLGLFLMTI
PWSPFNSALIRSPVIERKNAVFTNYVAASVTAISGSLAPROKISTMYVHSAVLA
GDVAVGTSCHLIPSPWLMAMVGLVAGLISVGAKXVLRGCNVLGIPIHSINGTNPSTL
LGLLEIYIVILVLDVTGAGNGMTGFQVLLSISGLSLATVIALISGLTLGLNLKTI
WKAPHAKKYFDQVFWKPRHLAVGF"

ORIGIN

Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGGGCTGTGTCGGCGCTGCTGCCCTTGAGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGGGCTGTGTCGGCGCTGCTGCCCTTGAGCCCTTAACACTG 60
QY 61 GAAGCAGCTCTGATCTCTCTCTTATTTTATTTTACCCCATAGAGGCTCTTATAGAGAT 120
Db 61 GAAGCAGCTCTGATCTCTCTCTTATTTTATTTTACCCCATAGAGGCTCTTATAGAGAT 120

QY 121 CAAAGGGGCTGAGGATCTTATCAAGTTGGCCAAAGATCTGACCGGATAGGGGCCATT 180
Db 121 CAAAGGGGCTGAGGATCTTATCAAGTTGGCCAAAGATCTGACCGGATAGGGGCCATT 180
QY 181 GGGTTGGGCTTCTCACCTGAGTTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
Db 181 GGGTTGGGCTTCTCACCTGAGTTTCCGAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
QY 241 CTCTTCATGCTGAGGCTTGTGTGTCAGTGGGCAATCTGTGTGACGGCTTCCGAGCCAG 300
Db 241 CTCTTCATGCTGAGGCTTGTGTGTCAGTGGGCAATCTGTGTGACGGCTTCCGAGCCAG 300
QY 301 TTCCCTTCTGGGAAGTGGTCAATCAACTGTTCAATTCGGGTGGCCACATAGATGCT 360
Db 301 TTCCCTTCTGGGAAGTGGTCAATCAACTGTTCAATTCGGGTGGCCACATAGATGCT 360
QY 361 TTGTCGGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCACTTGGCCGAGTTGGTGTG 420
Db 361 TTGTCGGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCACTTGGCCGAGTTGGTGTG 420
QY 421 GTGATGCTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCAGTAATATC 480
Db 421 GTGATGCTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCAGTAATATC 480
QY 481 TTCAACACAGACTACCAATGAATGATGCAATCTACGTTTGCAAGCTTATTTGGG 540
Db 481 TTCAACACAGACTACCAATGAATGATGCAATCTACGTTTGCAAGCTTATTTGGG 540
QY 541 CTGTCTGTGGCTGTGTGCTGCTGCAAGGCTTCAACCGAGGGAACGAGAGATTAAGATAG 600
Db 541 CTGTCTGTGGCTGTGTGCTGCTGCAAGGCTTCAACCGAGGGAACGAGAGATTAAGATAG 600
QY 601 ACAGCAACGATACCAAGTTTGTCTCCATGCTGGGCGCCTTCTTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCAAGTTTGTCTCCATGCTGGGCGCCTTCTTGTGATGTTCTGG 660
QY 661 CCAGTTTCAACTGCTCTGCTGAGAGTGCATTCGAAGGAAGAAATGCCGTTCATAC 720
Db 661 CCAGTTTCAACTGCTCTGCTGAGAGTGCATTCGAAGGAAGAAATGCCGTTCATAC 720
QY 721 ACCTCTATGCTGTAGCAGTACGCTGTGTGACAGCCATCTCAGGGTCACTCTTGCTCAC 780
Db 721 ACCTCTATGCTGTAGCAGTACGCTGTGTGACAGCCATCTCAGGGTCACTCTTGCTCAC 780
QY 781 CCCCAAGGGAAGATCAGCAACTTATGTGACAGTGCCTGGTGTGGCAGAGGCGTGGCT 840
Db 781 CCCCAAGGGAAGATCAGCAACTTATGTGACAGTGCCTGGTGTGGCAGAGGCGTGGCT 840
QY 841 GTGGGTACTGTGTACCTGATCCCTTCTCCGTGGCTTGCATGTGCTGGCTTGTG 900
Db 841 GTGGGTACTGTGTACCTGATCCCTTCTCCGTGGCTTGCATGTGCTGGCTTGTG 900
QY 901 GGTGGGCTGATCTCCGTCGGGGAGGCCAAGTACTCTGCGGGGTGTGTAAACCAAGTCTG 960
Db 901 GGTGGGCTGATCTCCGTCGGGGAGGCCAAGTACTCTGCGGGGTGTGTAAACCAAGTCTG 960
QY 961 GGAATCCCCACAGCTCATCATAGGAGTCAACTTCAAGCTTGGGTCTGCTTGAAGAG 1020
Db 961 GGAATCCCCACAGCTCATCATAGGAGTCAACTTCAAGCTTGGGTCTGCTTGAAGAG 1020
QY 1021 ATCATCTACATTGTGCTGTGTGTGTAACCGTCGAGCGGGCAATGGCAATGTGGC 1080
Db 1021 ATCATCTACATTGTGCTGTGTGTGTAACCGTCGAGCGGGCAATGGCAATGTGGC 1080
QY 1081 TTCAGAGTCTCTCTCAGCATATGGGCACTCAGCTTGCCATGTGATAGCTCTCAGCTCT 1140
Db 1081 TTCAGAGTCTCTCTCAGCATATGGGCACTCAGCTTGCCATGTGATAGCTCTCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTCATAGAGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGCTCTTAAATCTTAAATATGGAAGCACTCATAGAGCTTAA 1200

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 23:12:42 ; Search time 700 Seconds
(without alignments)
10604.794 Million cell updates/sec

Title: US-09-600-714-41

Sequence: 1 atgagctcaagtaaccgcg.....attggctgctgattttaa 1254

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	2	AAx86522
2	1252.4	99.9	1354	12	ADN88526
3	1252.4	99.9	1354	12	ADO47195
4	1252.4	99.9	2811	10	ADSO9785
5	1037	82.7	1810	13	ADRO7381
6	468	37.3	642	12	ADP28536
7	428.8	34.2	627	12	ADP28511
8	365.8	29.2	487	9	ACH43482
9	177	14.1	1805	6	ABK12333
10	175.4	14.0	1792	6	ABZ11480
11	175.4	14.0	1792	12	ADMA3998
12	165.4	13.2	1351	8	ACD20392
13	165.4	13.2	1853	3	AAC64286
14	165.4	13.2	1949	5	AAH25763
15	165.4	13.2	1952	6	ABK49215
16	165.4	13.2	1952	12	ADNO4909
17	165.4	13.2	1952	13	ADN82574
18	165.4	13.2	1968	8	ABX76257
19	165.4	13.2	1970	3	AAC76344
20	159.4	12.7	1679	6	ABK12334

C	21	152.4	12.2	3706	4	AAH18728	Aah18728 Human cDN
C	22	150.2	12.0	2098	6	ABK49216	Abk49216 cDNA enco
C	23	148.8	11.9	2415	10	ADB62576	Adb62576 Human cDN
C	24	147.8	11.8	339	10	ADSO7624	Adso7624 Novel cod
C	25	139.8	11.1	486	4	AAI11615	Aai11615 Probe #15
C	26	139.8	11.1	486	4	ABA53307	Abas3307 Human foe
C	27	139.8	11.1	486	4	AAI32911	Aai32911 Probe #15
C	28	139.8	11.1	486	4	ABA42886	Abas42886 Human bre
C	29	139.8	11.1	486	4	ABA23081	Abas23081 Probe #15
C	30	139.8	11.1	486	4	AAK27012	Aak27012 Human bon
C	31	139.8	11.1	486	4	AAK15568	Aak15568 Human bon
C	32	139.8	11.1	486	4	ABK26596	Abk26596 Human liv
C	33	139.8	11.1	486	5	AAI01542	Aai01542 Probe #15
C	34	139.8	11.1	486	6	ABSO1595	Abso1595 Human gen
C	35	135	10.8	1188	6	ADSO7444	Adso7444 Novel cod
C	36	127.2	10.1	954	12	ADR20082	Adr20082 Human imm
C	37	120	9.6	1571	6	ABL90445	Abi90445 Human pol
C	38	113.8	9.1	1840	4	ABL02133	Abi02133 Drosophill
C	39	111.8	8.9	123	4	AAI20829	Aai20829 Probe #10
C	40	111.8	8.9	123	4	ABA65900	Abas65900 Human foe
C	41	111.8	8.9	123	4	AAI46064	Aai46064 Probe #14
C	42	111.8	8.9	123	4	ABA48015	Abas48015 Human bre
C	43	111.8	8.9	123	4	ABA32986	Abas32986 Probe #11
C	44	111.8	8.9	123	4	AAK40045	Aak40045 Human bon
C	45	111.8	8.9	123	4	AAK14315	Aak14315 Human bra

ALIGNMENTS

RESULT 1	AAx86522	AAx86522 standard; cDNA; 1254 BP.
ID	XX	XX
AC	XX	AAx86522;
DT	XX	04-OCT-1999 (first entry)
DE	XX	cDNA sequence of the prevalent allele of the Rhd gene.
KW	XX	Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion; ss.
OS	XX	Homo sapiens.
FT	XX	Key
FT	XX	CDS
FT	XX	Location/Qualifiers
FT	XX	1..1254
FT	XX	/*tag= a
PN	XX	WO9937763-A2.
PD	XX	29-JUL-1999.
PF	XX	18-DEC-1998; 98WO-EP008319.
PR	XX	23-JAN-1998; 98BP-00101203.
PA	XX	(DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTTEMBERG.
PI	XX	Flegel WA, Wagner FF;
XX	XX	WPI; 1999-469127/39.
DR	XX	P-PADB; AAY24056.
PT	XX	Nucleic acid sequences correlated with Rhesus weak D phenotype. useful for screening blood from donors and recipients for transfusion methods.
XX	XX	Disclosure; Fig 2; 64pp; English.
XX	XX	The present sequence represents the prevalent allele of the Rhesus D (Rhd) antigen gene. The specification describes a Rhd contributing to or indicative of the weak D phenotype, where the Rhd polynucleotide carries at least one missense mutation as compared to the wild-type Rhd, in its transmembrane and/or intracellular regions, especially in amino acid

CC positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso that
 CC the D antigen does not carry a single missense mutation leading to a
 CC F223V or T263I substitution. The probes and antibodies are useful in the
 CC methods for detection of weak D phenotypes. Red blood cells, from
 CC probands, are useful for the assessment of the affinity, avidity and/or
 CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or
 CC of anti-globulin or anti-human-globulin antisera. Detecting the presence
 CC of the Rhd associated with weak D phenotype is useful for determining
 CC that a patient in need of a blood transfusion is to be transfused with
 CC Rhd negative blood from a donor. Alternatively, testing for weak D
 CC phenotype Rhd in the blood of a donor is useful for determining whether
 CC the donor blood should be excluded for transfusion to patients having
 CC wild type Rhd or weak D types, other than that of the donor weak D type
 CC XX
 SQ Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 U; 0 Other;

Query Match 100.0%; Score 1254; DB 2; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTTAACTACCCGCGCTGTGTCGCGGCTGCTGCCCCCTCTGCGGCCCTTAACACTG 60
 DB 1 ATGAGCTTAACTACCCGCGCTGTGTCGCGGCTGCTGCCCCCTCTGCGGCCCTTAACACTG 60
 QY 61 GAAGCAGCTCATCT 120
 DB 61 GAAGCAGCTCATCT 120
 QY 121 CAAAAGGGGCTGTGTCATCT 180
 DB 121 CAAAAGGGGCTGTGTCATCT 180
 QY 181 GGCCTTGGGCTTCT 240
 DB 181 GGCCTTGGGCTTCT 240
 QY 241 CTCTTCACTGCTGCGCTTGTGTCAGTGGGCAATCCGCTGCAAGCGCTTCTCTGAGCCAG 300
 DB 241 CTCTTCACTGCTGCGCTTGTGTCAGTGGGCAATCCGCTGCAAGCGCTTCTCTGAGCCAG 300
 QY 301 TTCCCTTCTGGGAAGGTGTGTCATCACTGTTCACTGTTGCGCTGCGCACTGAGTGTCT 360
 DB 301 TTCCCTTCTGGGAAGGTGTGTCATCACTGTTCACTGTTGCGCTGCGCACTGAGTGTCT 360
 QY 361 TTGCGGTGCTGATCTCACTGATGCTGTTGGGGAAGTCACTTGGCGCAGTTGGTG 420
 DB 361 TTGCGGTGCTGATCTCACTGATGCTGTTGGGGAAGTCACTTGGCGCAGTTGGTG 420
 QY 421 GTGATGCTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 480
 DB 421 GTGATGCTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 480
 QY 481 TTCAACACAGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 481 TTCAACACAGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 QY 541 CTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 600
 DB 541 CTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 600
 QY 601 ACAGGACAGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 DB 601 ACAGGACAGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 661 CCAAGTTTCACT 720
 DB 661 CCAAGTTTCACT 720
 QY 721 ACTTACTATGTGTGACAGTCAAGCTGTGTCAGCAGCATCTCAAGCTCACTTGGCTAC 780
 DB 721 ACTTACTATGTGTGACAGTCAAGCTGTGTCAGCAGCATCTCAAGCTCACTTGGCTAC 780
 QY 781 CCCCAAGGGAATGACAGACGACTTATGTGCACTGCGGTGTTGGCAGAGGCGTGTCT 840

DB 781 CCCCAAGGGAATGACAGACGACTTATGTGCACTGCGGTGTTGGCAGAGGCGTGTCT 840
 QY 841 GTGGGTACCTCGTGTCACTGATCCCTTCCCGGTGGCTTCCATGTGCTGGGTCTTGTG 900
 DB 841 GTGGGTACCTCGTGTCACTGATCCCTTCCCGGTGGCTTCCATGTGCTGGGTCTTGTG 900
 QY 901 GCTGGGTGATCTCTCGTGGGGGAGCCAAATGATCCTCCGAGGTGTGTAACGAGTCTG 960
 DB 901 GCTGGGTGATCTCTCGTGGGGGAGCCAAATGATCCTCCGAGGTGTGTAACGAGTCTG 960
 QY 961 GGGATTCGCCAGCTCCATCATAGGCTTCACTTACCTTGTGCTGGGTCTTGTGAGAG 1020
 DB 961 GGGATTCGCCAGCTCCATCATAGGCTTCACTTACCTTGTGCTGGGTCTTGTGAGAG 1020
 QY 1021 ATCATCTACATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1080
 DB 1021 ATCATCTACATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1080
 QY 1081 TTCCAGTCT 1140
 DB 1081 TTCCAGTCT 1140
 QY 1141 GGTCTCTGACAGGTTTGTCT 1200
 DB 1141 GGTCTCTGACAGGTTTGTCT 1200
 QY 1201 TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCTATTTGGCTGTGATTTTAA 1254
 DB 1201 TATTTTGAATGACCAAGTTTCTGGAAGTTTCTCTATTTGGCTGTGATTTTAA 1254

RESULT 2
 ADN88526
 ID ADN88526 standard; cDNA, 1354 BP.
 XX
 AC ADN88526;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human RHD gene cDNA sequence.
 XX
 KW mutation detection; exon nine; human; RHD; gene; Rh negative blood type;
 XX ss.
 OS Homo sapiens.
 XX
 PN JF2004089057-A.
 XX
 PD 25-MAR-2004.
 XX
 PF 30-AUG-2002; 2002JP-00254009.
 XX
 PR 30-AUG-2002; 2002JP-00254009.
 XX
 PA (NISE-) NIPPON SEKIJU GISHA.
 XX
 PA (WAKT) WAKUNAGA SEIYAKU KK.
 XX
 DR WPI; 2004-262639/25.
 XX
 PT Nucleic acid molecule useful for determining Rh negative blood type, has
 PT fragment which hybridizes to polynucleotide having mutation of guanine
 PT nucleotide to another nucleotide in 3' terminal of ninth exon.
 XX
 PS Claim 2; SEQ ID NO 1; 16pp; Japanese.
 XX
 CC The invention comprises a nucleic acid that can detect a mutation of a
 CC quantifiable residue to another nucleotide in the 3' terminal region of the
 CC ninth exon of the human RHD gene. The nucleic acid of the invention is
 CC useful for determining Rh negative blood type before birth and
 CC immediately after birth. The present cDNA sequence represents the human
 CC RHD gene.
 XX

SQ Sequence 1354 BP; 278 A; 346 C; 369 G; 361 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1252.4; DB 12; Length 1354;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG 60
DB 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG 60
QY 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTAAACCACTATGACGCTTCTTATGAGGAT 120
DB 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTAAACCACTATGACGCTTCTTATGAGGAT 120
QY 121 CAAAGGGGCTCGTGGGCTCTATCAAGTTGGCCAAAGCTGACCGGATGGCGGCATT 180
DB 121 CAAAGGGGCTCGTGGGCTCTATCAAGTTGGCCAAAGCTGACCGGATGGCGGCATT 180
QY 181 GGGTTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGCAGTGTGCGCTTCAC 240
DB 181 GGGTTGGGCTTCTCAGCTGAGTTTCCGAGACACAGCTGAGCAGTGTGCGCTTCAC 240
QY 241 CTCTTCACTGCTGGCGCTTGTGTGCAAGTGGCAATCTGCTGACGCGCTTCTGAGCCAG 300
DB 241 CTCTTCACTGCTGGCGCTTGTGTGCAAGTGGCAATCTGCTGACGCGCTTCTGAGCCAG 300
QY 301 TTGCGCTTCTGGGGAAGTGTGTCATCAGCTGTTCATGTTGGGCTGGCCACCATAGAGTCT 360
DB 301 TTGCGCTTCTGGGGAAGTGTGTCATCAGCTGTTCATGTTGGGCTGGCCACCATAGAGTCT 360
QY 361 TTGTGCGTGTGATCTTCAGTGTGATGCTGTCTTGGGGAAGTCACTTGGCGCAGTTGGTG 420
DB 361 TTGTGCGTGTGATCTTCAGTGTGATGCTGTCTTGGGGAAGTCACTTGGCGCAGTTGGTG 420
QY 421 GTGATGTGCTGTGTGAGGTGACAGCTTTTATGAGCACTGAGGATGTCATCAATATATC 480
DB 421 GTGATGTGCTGTGTGAGGTGACAGCTTTTATGAGCACTGAGGATGTCATCAATATATC 480
QY 481 TTGACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTGACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGTGCTGTGTGCTGTGCTGCAAAAGCTCTTACCCGAGGGAACGAGATTAAGATCAG 600
DB 541 CTGTCTGTGCTGTGTGCTGTGCTGCAAAAGCTCTTACCCGAGGGAACGAGATTAAGATCAG 600
QY 601 ACAGCAACGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ACAGCAACGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 CCAAGTTTCAACTCTGCTGTGAGAAATCAATCAAAAGAAATGCGGTGTTAAC 720
DB 661 CCAAGTTTCAACTCTGCTGTGAGAAATCAATCAAAAGAAATGCGGTGTTAAC 720
QY 721 ACCTACTATCTGTAGAGCTGAGCTGATGACAGCCATCTCAGGGTATCTTGGCTCAC 780
DB 721 ACCTACTATCTGTAGAGCTGAGCTGATGACAGCCATCTCAGGGTATCTTGGCTCAC 780
QY 781 CCCCAAGGGAAGATCAGCAAGACTTATATGTGACAGTCCGCTGTTGCAAGAGCGTGGCT 840
DB 781 CCCCAAGGGAAGATCAGCAAGACTTATATGTGACAGTCCGCTGTTGCAAGAGCGTGGCT 840
QY 841 GTGGGTATCTGTGTGACAGTATCCCTTCCGCGGTGTCGATGAGTGTGGCTTGTG 900
DB 841 GTGGGTATCTGTGTGACAGTATCCCTTCCGCGGTGTCGATGAGTGTGGCTTGTG 900
QY 901 GCTGGGCTGATCTCTGCTGTGGGGAAGCAAGTACTGCGGGGTGTTTAAACGAGTCTG 960
DB 901 GCTGGGCTGATCTCTGCTGTGGGGAAGCAAGTACTGCGGGGTGTTTAAACGAGTCTG 960
QY 961 GGGATTTCCCAACAGCTTCATCATGAGGCTACAACTTCAAGCTTGTGGGTCTGTTGAGAG 1020
DB 961 GGGATTTCCCAACAGCTTCATCATGAGGCTACAACTTCAAGCTTGTGGGTCTGTTGAGAG 1020
  
```

QY 1021 ATCATCTACATTTGTGCTGTGTGCTGTGATACCGTGGAGCCGCAATGGCATGATTGGC 1080
 DB 1021 ATCATCTACATTTGTGCTGTGTGCTGTGATACCGTGGAGCCGCAATGGCATGATTGGC 1080
 QY 1081 TTCCAGGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATTCGTGATAGCTTCAAGTCT 1140
 DB 1081 TTCCAGGTCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCCATTCGTGATAGCTTCAAGTCT 1140
 QY 1141 GGTCTCTGACAGTTTGTCTCTTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 DB 1141 GGTCTCTGACAGTTTGTCTCTTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 QY 1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGGCTGTGATTTTAA 1254
 DB 1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGGCTGTGATTTTAA 1254

RESULT 3
 ADO47195
 ID ADO47195 standard; cDNA, 1354 BP.
 AC ADO47195;
 AC 15-JUL-2004 (first entry)
 DT 15-JUL-2004 (first entry)
 DE cDNA sequence of human Rhd blood group antigen.
 KM immunoglobulin, B cell; germline; region V; region D; region J;
 KM recombination-promoting protein; recombination activating gene 1; RAG-1;
 KM RAG-2; ss; terminal deoxynucleotidyltransferase; Tdt; V(D)J recombinase;
 KM Rhd blood group antigen.
 OS Homo sapiens.
 XX
 XX WO2004029249-A1.
 XX
 XX 08-APR-2004.
 XX
 XX 30-SEP-2003; 2003WO-AU001286.
 XX
 XX 30-SEP-2002; 2002US-0415024P.
 XX
 XX (CENT-) CENTENARY INST CANCER MEDICINE & CELL BI.
 XX
 XX Jolly C;
 XX
 XX WPI; 2004-316126/29.
 XX
 XX
 XX New vector comprising one or more immunoglobulin regions selected from V,
 PT D, and J regions, a 5' flanking region, and a 3' flanking region, useful
 PT for reverting cell lines to a pro-B cell-like state or to a germline-like
 PT state.
 XX
 XX
 XX Example 4; Page 322-323; 382pp; English.
 XX
 XX The specification describes a method for the preparation of
 CC immunoglobulins. The method uses a vector for reverting cell lines to a
 CC pro-B cell-like state or to a germline-like state, by replacing one or
 CC more of the immunoglobulin regions V, D, and J of the rearranged
 CC immunoglobulin gene with the germ-line or unrearranged V, D or J regions.
 CC The vector can comprise a polynucleotide sequence encoding a
 CC recombination-promoting protein, selected from recombination activating
 CC gene 1 (RAG-1), RAG-2, terminal deoxynucleotidyltransferase (Tdt). These
 CC proteins collectively constitute a V(D)J recombinase. The method is
 CC useful for the preparation of immunoglobulin heavy and/or light chains
 CC and their fragments. The present sequence encodes a human Rhd blood group
 CC antigen, and is in the course of the invention.
 XX
 XX
 SQ Sequence 1354 BP; 278 A; 346 C; 369 G; 361 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1252.4; DB 12; Length 1354;
 Best Local Similarity 99.9%; Pred. No. 0;

Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCAGCGGATGTCGTCGGCGGTGCTGCTCCCTCTGGGCGCTTAACATCG 60
Db 1 ATGAGCTCTAAGTACCAGCGGATGTCGTCGGCGGTGCTGCTCCCTCTGGGCGCTTAACATCG 60
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTTTACCAGTATGACGCTTCTTAAGAGAT 120
Db 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTTTACCAGTATGACGCTTCTTAAGAGAT 120
QY 121 CAAAAGGGGCTGTGGCATCTTCAAGTTGGGCAAGATCTGACCGTGATGGCGCCATT 180
Db 121 CAAAAGGGGCTGTGGCATCTTCAAGTTGGGCAAGATCTGACCGTGATGGCGCCATT 180
QY 181 GGGTTGGGCTTCTCTCACTTGAGTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAAC 240
Db 181 GGGTTGGGCTTCTCTCACTTGAGTTCCGGAGACACAGCTGGAGCAGTGGGCTTCAAC 240
QY 241 CTCTTCATGCTGCGCTTGGTGTGCAAGTGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
Db 241 CTCTTCATGCTGCGCTTGGTGTGCAAGTGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
QY 301 TTCCCTTCTGGAGAGTGGTCACTCACTGTTCACTATTCGGCTGGCCACCATGAGTCT 360
Db 301 TTCCCTTCTGGAGAGTGGTCACTCACTGTTCACTATTCGGCTGGCCACCATGAGTCT 360
QY 361 TTGTGCTGCTGATCTCAGTGGATCTGTCTTGGGGAAGTCACTTGGCGCAGTTGGT 420
Db 361 TTGTGCTGCTGATCTCAGTGGATCTGTCTTGGGGAAGTCACTTGGCGCAGTTGGT 420
QY 421 GTGATGCTGCTGAGTGAAGTGAAGCTTAAAGCACTGAGAGTGTGATGATATATC 480
Db 421 GTGATGCTGCTGAGTGAAGTGAAGCTTAAAGCACTGAGAGTGTGATGATATATC 480
QY 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 ACAGCAACGATACCAGTTGTCTGCGCATGCTGCGCGCCCTCTTCTGTGAGATGTTCTG 660
Db 601 ACAGCAACGATACCAGTTGTCTGCGCATGCTGCGCGCCCTCTTCTGTGAGATGTTCTG 660
QY 661 CCAAGTTTCAACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CCAAGTTTCAACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ACCCTATGCTGCTGAGAGTCAAGGTCAGGCACTTCAAGGTCATCTTGGCTCAC 780
Db 721 ACCCTATGCTGCTGAGAGTCAAGGTCAGGCACTTCAAGGTCATCTTGGCTCAC 780
QY 781 CCCCAGGGAAGATCAGCAAGACTTATGTCACAGTGGCTGTTGGCAGAGCGGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGACTTATGTCACAGTGGCTGTTGGCAGAGCGGCT 840
QY 841 GTGGGATACCTCGGTGACCTGATCCCTTCTCCGTGGCTTGCATGGTGGCTCTTGTG 900
Db 841 GTGGGATACCTCGGTGACCTGATCCCTTCTCCGTGGCTTGCATGGTGGCTCTTGTG 900
QY 901 GCTGGGCTGATCCGCTGGGGAGGCAAGTACCTGCGGGGTTGTTAAGCAAGTGTG 960
Db 901 GCTGGGCTGATCCGCTGGGGAGGCAAGTACCTGCGGGGTTGTTAAGCAAGTGTG 960
QY 961 GGGATTTCCCAACAGCTCCATCATGGGCTTAACTTCAAGCTTCTGGTCTGTTGAGAG 1020
Db 961 GGGATTTCCCAACAGCTCCATCATGGGCTTAACTTCAAGCTTCTGGTCTGTTGAGAG 1020
QY 1021 ATCATCTCATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 ATCATCTCATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

```

```

QY 1081 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGGATAGCTCAGCTCT 1140
Db 1081 TTCCAGGTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGGATAGCTCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAAATAATGAAACACCTCATGAGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAAATAATGAAACACCTCATGAGCTTAA 1200
QY 1201 TATTTGATACCAAGTTTCTGAAATTTCCCATTTGGCTGTGGATTTTAA 1254
Db 1201 TATTTGATACCAAGTTTCTGAAATTTCCCATTTGGCTGTGGATTTTAA 1254

```

RESULT 4
ADE09785
ADE09785 standard; DNA; 2811 BP.

AC ADE09785;

DT 29-JAN-2004 (first entry)

DE Novel DNA-related contig nucleotide sequence #507.

KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig; ds.

OS Unidentified.

PN MO2003054152-A2.

PD 03-JUL-2003.

PF 10-DEC-2002; 2002MO-US039555.

PR 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

PA (HYSB-) HYSB INC.

PI Tang YF, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

PI Ghosh M, Xue AU, Wehman T, Weng G, Zhou P, Drmanac RT, Wang Z;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI WPI; 2003-569235/53.

PS New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

PS Disclosure; SEQ ID NO 2329; 1177bp; English.

CC The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to

CC compare with endogenous DNA sequences in patients to identify potential

CC genetic disorders. The present DNA sequence was used in the

CC exemplification of the invention.

SO Sequence 2811 BP; 658 A; 690 C; 659 G; 804 T; 0 U; 0 Other;

Query Match 99.9%; Score 1252.4; DB 10; Length 2811;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	TTGAGCTCTAAGTACCCGGGGCTGTCCGGGGCTGCGCCCTCTGGGGCCCTAAACCTG	60
Db	17	ATGAGCTCTAAGTACCCGGGGCTGTCCGGGGCTGCGCCCTCTGGGGCCCTAAACCTG	76
Qy	61	GAACAGCTCTCATTTCTCTCTTCTCTAATTTTTCACCTATGACAGCTTCTTATGAGAGAT	120
Db	77	GAACAGCTCTCATTTCTCTCTTCTCTAATTTTTCACCTATGACAGCTTCTTATGAGAGAT	136
Qy	121	CAAAAGGGGCTCGTGGCATTCCTATCAAGTTGGCCAAATCTGACCGGTATGGCGGCATT	180
Db	137	CAAAAGGGGCTCGTGGCATTCCTATCAAGTTGGCCAAATCTGACCGGTATGGCGGCATT	196
Qy	181	GGCTTGGGCTTCTTCACCTCGAGTTTCCGAGACACACAGCTGGAGCAGTGTGGCTTCAAC	240
Db	197	GGCTTGGGCTTCTTCACCTCGAGTTTCCGAGACACACAGCTGGAGCAGTGTGGCTTCAAC	256
Qy	241	CTCTTCATGCTGGGCGCTTGATGTGACAGGGGCAATCCGTCTGGACCGGCTTCTGAGCCAG	300
Db	257	CTCTTCATGCTGGGCGCTTGATGTGACAGGGGCAATCCGTCTGGAGCGGCTTCTTGAAGCCAG	316
Qy	301	TTCCCTTCTGGGAAAGTGGTCATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGT	360
Db	317	TTCCCTTCTGGGAAAGTGGTCATCACACTGTTCAAGTATTCGGCTGGCCACCATGAGTGT	376
Qy	361	TTGTTCGGTCTGATCTCAGTGGATGTCTCTTGGGAAAGTCAACTTGGCCAGTTGGTG	420
Db	377	TTGTTCGGTCTGATCTCAGTGGATGTCTCTTGGGAAAGTCAACTTGGCCAGTTGGTG	436
Qy	421	GTGATGGGCTGGTGGAGGTGACACAGCTTTAGGCAACCTGAGGATGGTCATGATATATC	480
Db	437	GTGATGGGCTGGTGGAGGTGACACAGCTTTAGGCAACCTGAGGATGGTCATGATATATC	496
Qy	481	TTCAACACAGACTACCAATGAAACATGATGACATCTACGTTCGACGCTATTTTGGG	540
Db	497	TTCAACACAGACTACCAATGAAACATGATGACATCTACGTTCGACGCTATTTTGGG	556
Qy	541	CTGTCTGTGGCTGTGGCTGCTCCGCAAGCTCTTACCCGAGGAAAGGAGATTAAGATCAG	600
Db	557	CTGTCTGTGGCTGTGGCTGCTCCGCAAGCTCTTACCCGAGGAAAGGAGATTAAGATCAG	616
Qy	601	ACAGCAACGATACCAGTTGTGTGCGCATGCTGGGGCGGCTCTTCTTGATGATGTTCTGG	660
Db	617	ACAGCAACGATACCAGTTGTGTGCGCATGCTGGGGCGGCTCTTCTTGATGATGTTCTGG	676
Qy	661	CCAAGTTTCAACTCTGCTCTGCTGAGAAAGTCCAATCGAAAGAAAGATCCGCTTTCAC	720
Db	677	CCAAGTTTCAACTCTGCTCTGCTGAGAAAGTCCAATCGAAAGAAAGATCCGCTTTCAC	736
Qy	721	ACCTACTATGCTGTAGCAGTCAGCGTGTGTACAGCCATCTCAGGGTCACTCTTGGCTCAC	780
Db	737	ACCTACTATGCTGTAGCAGTCAGCGTGTGTACAGCCATCTCAGGGTCACTCTTGGCTCAC	796
Qy	781	CCCCAAGGGAAGATCAGCATACTTATGTGACACAGTGGGTGTGGCAGAGGCGTGGCT	840
Db	797	CCCCAAGGGAAGATCAGCATACTTATGTGACACAGTGGGTGTGGCAGAGGCGTGGCT	856
Qy	841	GTGGGTACTCTGTGTCACTGATCCCTTCTCCGTGGCTTCCGATGATGATGATGATGATG	900
Db	857	GTGGGTACTCTGTGTCACTGATCCCTTCTCCGTGGCTTCCGATGATGATGATGATGATG	916
Qy	901	GCTGGGCTGATCTCCGTGGGGGAGAACCAAGTACCTGCCGGGGGTGTGTAAACGAGTGTG	960
Db	917	GCTGGGCTGATCTCCGTGGGGGAGAACCAAGTACCTGCCGGGGGTGTGTAAACGAGTGTG	976
Qy	961	GGGATTTCCCAACAGCTTCATCATGGGCTAACCACTTCACTTGGGTGTGGTGTGGAGAG	1020
Db	977	GGGATTTCCCAACAGCTTCATCATGGGCTAACCACTTCACTTGGGTGTGGTGTGGAGAG	1036
Qy	1021	ATCATCTACATTTGTGCTGTGGTGTATATCCGTGGAGCGCGGCAATGGGATGATTTGGC	1080
Db	1037	ATCATCTACATTTGTGCTGTGGTGTATATCCGTGGAGCGCGGCAATGGGATGATTTGGC	1096
Qy	1081	TTTCCAGGCTCTTCCAGCAATTGGGGAACCTGAGCTTGGCCATCTGATGATGATCTCAGCTCT	1140

Db	1097	TTCCAGGCTCTTCCAGCATGGGGAACCTAGCTTGCCATCGATAGCTCTTCAGGCT	1156
Qy	1141	GGTCTCCGTGACAGGTTTGCTCTTAATCTTAAATATGAAAGCACCCTCATGAGCTTAA	1200
Db	1157	GGTCTCTCGACAGGTTTGCTCTTAAATCTTAAATATGAAAGCACCCTCATGAGCTTAA	1216
Qy	1201	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA	1254
Db	1217	TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA	1270
RESULT 5			
ID	ADR07381	standard; cDNA; 1810 BP.	
AC	ADR07381;		
XX			
DT	04-NOV-2004	(first entry)	
XX			
DE	Full length human cDNA useful for treating neurological disease Seq 887.		
XX			
KM	gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;		
KM	osteoporosis; neurological disease; Alzheimer's disease;		
KM	Parkinson's disease; dementia; short memory; cancer;		
KM	sense or motor function; emotional reaction; fear response; panic;		
KM	osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;		
XX	tranquilliser.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1447413-A2.		
XX			
PD	18-AUG-2004.		
XX			
PF	12-FEB-2004; 2004EP-00003145.		
XX			
PR	14-FEB-2003; 2003JP-00102207.		
PR	09-MAY-2003; 2003JP-00131452.		
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX			
PI	Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;		
PI	Wakamatsu A, Ishii S, Nagai K, Irie R;		
XX			
DR	WPI: 2004-583265/57.		
XX	P-ESDB; ADR09337.		
PT	New 1995 cDNA, useful for treating osteoporosis, neurological diseases,		
PT	Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.		
PS	Claim 1; SEQ ID NO 887; 2686pp; English.		
CC	This invention relates to novel, isolated full length human cDNA		
CC	molecules and the encoded proteins thereof. Specifically, it refers to		
CC	cDNA clones obtained by an oligo-capping method, where none of these		
CC	clones are identical to any known human mRNAs. The present invention		
CC	describes an immunoassay to identify agonists and antagonists, as well as		
CC	antibodies, antisense molecules and siRNAs that can all be used to bind		
CC	to and modulate expression of the cDNA molecules. As such, these		
CC	molecules are useful for diagnostic markers or therapeutic targets for		
CC	the various diseases or morbid states. In particular, they are useful in		
CC	gene therapy for treating osteoporosis, neurological disease, Alzheimer's		
CC	disease, Parkinson's disease, dementia, short memory and various cancers,		
CC	as well as for maintaining equilibrium of sense or motor function, and		
CC	for treating emotional reaction, fear response and panic. Accordingly,		
CC	they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,		
CC	cytostatic and tranquilliser activities. This polynucleotide is a full		
CC	length human cDNA sequence of the invention. NOTE: This sequence is not		
CC	CD-ROM from the European Patent Office, Vienna Sub-office.		
XX	Sequence 1810 BP; 386 A; 464 C; 504 G; 456 T; 0 U; 0 Other;		

Query Match 82.7%; Score 1037; DB 13; Length 1810;
Best Local Similarity 93.9%; Pred. No. 2,7e-289;
Matches 1079; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
QY 106 GCTTCCTTAAGAGATCAAAAGGGGCTCGTGGCACTCTATCAAGTTGGCCAAAGATCGACC 165
DB 447 GCTTCCTTAAGATCAAAAGGGGCTGGAGCACTGGCGGCTGGCCAAAGATCGACC 506
QY 166 GTGATGGCGGCAATGGGCTTCTCTCACTCGAGTTCCGAGAGACAGCTGGAGC 225
DB 507 GTGATGGCGGCGCTTGGCTTGGCTTCTCTCACTCAAAATTCGAGAGACAGCTGGAGC 566
QY 226 AGTGTGGCTTCAACTCTTCACTGCTGGCGCTTGTGTGACAGTGGGCAATCTGCTGGAC 285
DB 567 AGTGTGGCTTCAAACTCTTCACTGCTGGCGCTTGTGTGACAGTGGGCAATCTGCTGGAC 626
QY 286 GGGTCTTGAGGCAAGTTCCCTCTGGAGAGTGTGATCACTGTTCAATTTGAGTGGCTG 345
DB 627 GGGTCTTGAGGCAAGTTCCCTCTGGAGAGTGTGATCACTGTTCAATTTGAGTGGCTG 686
QY 346 GGCACCATGAGTCTTGTGTGTGTGATCTGATGATGATGATGATGATGATGATGATGATG 405
DB 687 GGCACCATGAGTCTTGTGTGTGTGATCTGATGATGATGATGATGATGATGATGATGATG 746
QY 406 TTGGGCGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 465
DB 747 TTGGGCGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 806
QY 466 GTCATCAGTAATATCTTCAACACAGACTACCAATGACATGATGACATGATGATGATGATG 525
DB 807 GTCATCAGTAATATCTTCAACACAGACTACCAATGACATGATGACATGATGATGATGATG 866
QY 526 GCAGCTTATTTTGGGCTGTCTGTGGCTGTGTGGCTGTGGCCAAAGCCTTACCCAGGAGAC 585
DB 867 GCAGCTTATTTTGGGCTGTCTGTGGCTGTGTGGCTGTGGCCAAAGCCTTACCCAGGAGAC 926
QY 586 GAGGATTAAGATACAGACAGCAACGATACCCAGTTTGTGTGTGTGTGTGTGTGTGTGTGT 645
DB 927 GAGGATTAATGATCAAGAGACAGCAACGATACCCAGTTTGTGTGTGTGTGTGTGTGTGTGT 986
QY 646 TTGTGATGTTTGTGGCCAAAGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
DB 987 TTGTGATGTTTGTGGCCAAAGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1046
QY 706 AATGCCGTTTCAACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
DB 1047 AATGCCATGTTTCAACACTATGATGATGATGATGATGATGATGATGATGATGATGATG 1106
QY 766 TCATCTTGGCTCACTCCCAAGGAGATGACAGCACTTATGATGATGATGATGATGATGATG 825
DB 1107 TCATCTTGGCTCACTCCCAAGGAGATGACAGCACTTATGATGATGATGATGATGATGATG 1166
QY 826 GCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 885
DB 1167 GCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1226
QY 886 GTGTGGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
DB 1227 GTGTGGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1286
QY 946 TGTAAACGAGTGTGGGAGTTCCCAAGTCCATCATGGGCTACAACTTACAGTTGTGCTG 1005
DB 1287 TGTAAACGAGTGTGGGAGTTCCCAAGTCCATCATGGGCTACAACTTACAGTTGTGCTG 1346
QY 1006 GGTCTCTTGAGAGATCACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1065
DB 1347 GGTCTCTTGAGAGATCACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1406
QY 1066 AATGCGATATTTGCTTCAAGTCTCTCAAGATTTGGGAGACTCAGCTGGGCAATGCTGTG 1125
DB 1407 AATGCGATATTTGCTTCAAGTCTCTCAAGATTTGGGAGACTCAGCTGGGCAATGCTGTG 1466
```

```
QY 1126 ATAGCTCTACGTCGTCTCTCTGACAGGTTTGTCTCTTAATCTTAATATGAAAGCA 1185
DB 1467 ATAGCTCTACGTCGTCTCTCTGACAGGTTTGTCTCTTAATCTTAATATGAAAGCA 1526
QY 1186 CCTCATGAGGCTAAATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTT 1245
DB 1527 CCTCATGAGGCTAAATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTT 1586
QY 1246 GGATTTTAA 1254
DB 1587 GGATTTTAA 1595

RESULT 6
ADP28536
ID ADP28536 standard; DNA; 642 BP.
AC ADP28536;
XX
XX
DT 12-AUG-2004 (first entry)
DE Human secreted protein encoding sequence SEQ ID #534.
XX
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; ds; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
```


PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-041101P.
PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 18-APR-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;
PI Helebeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PI genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 534; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein encoding sequence. The
CC present sequence is available on WIPOWEB and is not in the specification.

XX Sequence 642 BP; 112 A; 163 C; 196 G; 171 T; 0 U; 0 Other;

Query Match 37.3%; Score 468; DB 12; Length 642;

Best Local Similarity 81.0%; Pred. No. 1.1e-124;
Matches 632; Conservative 0; Mismatches 10; Indels 138; Gaps 2;

QY 154 CAAGATCTGACCGTGATGGCGCCATTGGCTTGGCTTCTCACTCGAGTTTCCGAGA 213
DB 1 CAAGATCTGACCGTGATGGCGCCATTGGCTTGGCTTCTCACTCGAGTTTCCGAGA 60
QY 214 CACAGCTGAGAGCACTGTGGCTTCAACCTCTTCACTGTGGCGCTTGGTGTGCACTGGCA 273
DB 61 CACAGCTGAGAGCACTGTGGCTTCAACCTCTTCACTGTGGCGCTTGGTGTGCACTGGCA 120
QY 274 ATCTGTGAGACGGCTTCTGAGCAGATTCCTCTGAGAAAGTGTGATCAGACTGTTC 333
DB 121 ATCTGTGAGACGGCTTCTGAGCAGATTCCTCTGAGAAAGTGTGATCAGACTGTTC 180
QY 334 AGTAATGGCTGGCCACCATGAGTCTTGTGGTGTGATCTGAGTGTCTGTTC 393

DB 181 AGTAATGGCTGGCCACCATGAGTCTTGTGGTGTGATCTGAGTGTCTGTTC 240
QY 394 GGAAGGTCAACTTGGCGCAGTTGTGTGATGTGTGTGAGTGTGAGCTTTAGGC 453
DB 241 GGAAGGTCAACTTGGCGCAGTTGTGTGATGTGTGTGAGTGTGAGCTTTAGGC 300
QY 454 AACCTGAGATGTGATCACTAATATCTTCAACACAGACTTACCATGAAATGATGAC 513
DB 301 AACCTGAGG-TGGTCACTACGATTAATCTTCAA----- 331
QY 514 ATCTAGTGTTCGACGCTATTTTGGGCTGTGTGGCTGTGGCTGCCAAAGCTCTTA 573
DB 332 ----- 331
QY 574 CCCGAGGAACGAGAGATAAAGATCAGACAGACATACCAGTTTGTCTGCATCTG 633
DB 332 -----CTGCCATGCTG 342
QY 634 GGGGCCCTCTTCTGTGTGATGTTCTGGCCAAGTTTCACTGTCTGTGAGAGTCCA 693
DB 343 GGGGCCCTCTTCTGTGTGATGTTCTGGCCAAGTGTCACTCTCTGTGTGAGAGTCCA 402
QY 694 ATGGAAGGAAGAATGCCGTGTTCAACACTACTATGTCTAGACAGTCACTGTGTGACA 753
DB 403 ATCCAAAGGAAGATGACATGTTCAACACTACTATGTCTAGACAGTGTGTGACA 462
QY 754 GCCATCTCAGAGTCACTCTTGGCTCAACCCCAAGGAAGATCAGCAAGCTTATGTGCAC 813
DB 463 GCCATCTCAGAGTCACTCTTGGCTCAACCCCAAGGAAGATCAGCAAGCTTATGTGCAC 522
QY 814 AGTCCGCTGTGGCAGAGAGCGTGGCTGTGGGTAACCTGTGTCACTGTATCCCTTC 873
DB 523 AGTCCGCTGTGGCAGAGAGCGTGGCTGTGGGTAACCTGTGTCACTGTATCCCTTC 582
QY 874 TGGCTTGCCTCACTGT 933
DB 583 TGGCTTGCCTCACTGT 642

RESULT 7

ID ADP28511 standard; DNA; 627 BP.

XX ADP28511;

DT 12-AUG-2004 (first entry)

XX Human secreted protein encoding sequence SEQ ID #509.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; der; human secreted protein.

XX Homo sapiens.

PN WO2004035732-A2.

PD 29-Apr-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411075P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411011P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463712P.
PR 02-MAY-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 08-JUL-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,
PI Halebek R, Huang MM, Kochakota S, Haisan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 509; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein encoding sequence. The
CC present sequence is available on WIPWEB and is not in the specification.
XX

SO Sequence 627 BP; 119 A; 171 C; 180 G; 157 T; 0 U; 0 Other;

Query Match 34.2%; Score 428.8; DB 12; Length 627;
Best Local Similarity 77.6%; Pred. No. 2.4e-113;
Matches 605; Conservative 0; Mismatches 22; Indels 153; Gaps 1;

QY 154 CAGATCTGACCGTATGCGGCGCATTTGGGCTTCTCACTGATTCGCGAGA 213
Db 1 CAGATCTGACCGTATGCGGCGCATTTGGGCTTCTCACTGATTCGCGAGA 60
QY 214 CACAGCTGAGGAGTGTGGCTTCAACCTTTATGCTGGCGCTTGTGAGTGGCA 273
Db 61 CACAGCTGAGGAGTGTGGCTTCAACCTTTATGCTGGCGCTTGTGAGTGGCA 120
QY 274 ATCTGCTGAGCGGCTTCTGAGCCAGTTCCCTTGGGAGTGTATCACTGCTTC 333
Db 121 ATCTGCTGAGCGGCTTCTGAGCCAGTTCCCTTGGGAGTGTATCACTGCTTC 180
QY 334 AGTATTCGGCTGGCCACCATGAGTGTCTTGTGCTGATCTAGTGAATGCTTGTG 393
Db 181 AG-----
QY 394 GGGAGATGACCTTGGCGGAGTGTGATGATGCTGTGGAGTACACCTTTAGGC 453
Db 183 -----
QY 454 AACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 513
Db 183 -----ACAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
QY 514 ATCTAGTGTGCGACGCTATTTTGGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGT 573
Db 208 TTCTAGTGTGCGACGCTATTTTGGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGT 267
QY 574 CCGAGGGAAGGAGTAAAGATCAAGACAGACATACAGTGTGTGTGTGTGTGTGTGT 633
Db 268 CCGAGGGAAGGAGTAAAGATCAAGACAGACATACAGTGTGTGTGTGTGTGTGTGT 327
QY 634 GCGGCGCTTCTTGT 693
Db 328 GCGGCGCTTCTTGT 387
QY 694 ATGAAAGGAAGATGCGGTTCACACCTATCTATCTATCTATCTATCTATCTATCTAT 753
Db 388 ATGAAAGGAAGATGCGGTTCACACCTATCTATCTATCTATCTATCTATCTATCTAT 447
QY 754 GGCATCTCAGGCTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 813
Db 448 GGCATCTCAGGCTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 507
QY 814 AGTCCGCTGTGCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 873
Db 508 AGTCCGCTGTGCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567
QY 874 TGGCTTGCATGT 933
Db 568 TGGCTTGCATGT 627

RESULT 8
ACH43482
ID ACH43482 standard; cDNA, 487 BP.
XX
AC ACH43482;
XX
DT 13-OCT-2003 (first entry)
XX

DE	Human foetal brain cDNA #4207.
XX	
KM	Human; 88; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW	genome mapping; biodiversity; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	US2003073623-A1.
PD	
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI; 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
PS	
XX	Claim 1; SEQ ID NO 30694; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=20030073623
XX	
SQ	Sequence 487 BP; 86 A; 120 C; 156 G; 119 T; 0 U; 6 Other;
XX	
Query Match	29.2%; Score 365.8; DB 9; Length 487;
Beet Local Similarity	99.5%; Pred. No. 3.8e-95;
Matches 367; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	706 AATGCCGCTGTTCACACCTACTATGCTTAGACGTCAGCGTGATGACAGCATCTCAGGG 765
DB	56 ATTCOCGGTTCACACCTACTACTGCTGTGAAGATCAAGCGTGATGACAGCATCTCAGGG 115
QY	766 TCATCTCTTGAGCTCACCCCAAGAGGAGATCAGCAAGACTTATGTGCACAGTCCGGTGTG 825
DB	116 TCATCTCTTGAGCTCACCCCAAGAGGAGATCAGCAAGACTTATGTGTGCACAGTCCGGTGTG 175
QY	826 GCAGGAGCGTGAGCTGTGGATACCTCGATGACCTGATCCCTTCCGATGGCTTGCAGT 885
DB	176 GCAGGAGCGTGAGCTGTGGATACCTCGATGACCTGATCCCTTCCGATGGCTTGCAGT 235
QY	886 GTGCTGGCTCTTGTGGCTGTGGCTGATCTCCGTCCGGGGAGCCAAATACCTGCCGGGGTGT 945
DB	236 GTGCTGGCTCTTGTGGCTGTGGCTGATCTCCGTCCGGGGAGCCAAATACCTGCCGGGGTGT 295
QY	946 TGTAAACGAGTGTGGGATTTCCCAACAGCTTCATCATGAGCTTCAACTTACGTTTGCT 1005

Db	296	TGTAAACGAGTCTGGGGATTCCCCACAGCTCCATCATGGGCTACAACTTCAGCTTGCTG	355
Oy	1006	GGTCTGCTTGAGAGATCATCTAATATTGCTGCTGCTGCTTGAATACCGTCGAGCCGGC	1065
Db	356	GGTCTGCTTGAGAGATCATCTAATATTGCTGCTGCTGCTTGAATACCGTCGAGCCGGC	415
Oy	1066	AATGGCATTG 1074	
Db	416	AATGGCATTG 424	
RESULT 9			
ABK12333	1D	ABK12333 standard; cDNA; 1805 BP.	
XX	AC	ABK12333;	
XX	DT	05-JUN-2002 (first entry)	
XX	DE	cDNA encoding human nonerythroid Rh glycoprotein RhBG.	
XX	KW	Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;	
KW	KW	Rh type B transporter gene; polypeptic transporter-type protein;	
KW	KW	ion transporter; chromosome 1q21.3; gene; sb.	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
XX	FD	39..1415	
FD	FD	/*tag= a	
FD	FD	/product= "Human RhBG protein"	
FD	FD	1248..1377	
FD	FD	/*tag= b	
FD	FD	/note= "Encodes C-tail"	
FD	FD	1769..1774	
FD	FD	/*tag= C	
FD	FD	/standard_name= "PolyA signal"	
FD	FD	/note= "Atypical polyadenylation site"	
XX	XX	WO200216396-A2.	
PD	PD	28-FEB-2002.	
XX	XX	17-AUG-2001; 2001WO-US025881.	
XX	XX	21-AUG-2000; 2000US-0226767P.	
XX	XX	(NYBL-) NEW YORK BLOOD CENT INC.	
PA	PA	Huang C, Liu Z;	
XX	PI	WPI; 2002-280905/32.	
XX	DR	P-PSDB; AAU78091.	
XX	XX	Homologs of mouse or human nonerythroid Rh glycoproteins, Rhbg or RhBG,	
XX	PT	respectively, useful in the production of antibodies which are useful for	
PT	PT	detecting Rhbg or RhBG glycoproteins in a sample.	
XX	XX	Claim 1; Fig 1a; 59pp; English.	
XX	XX	The present invention relates to a new protein or peptide comprising an	
CC	CC	amino acid sequence having at least 60% identity to a sequence comprising	
CC	CC	455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human	
CC	CC	nonerythroid Rh glycoprotein homologue (RhBG)) amino acids, fully defined	
CC	CC	in the specification. The antibody of the invention is useful for	
CC	CC	detecting an Rhbg or an RhBG glycoprotein in a sample, by contacting the	
CC	CC	sample with antibody under conditions suitable for binding, assessing the	
CC	CC	specific binding to the antibody, and thus detecting the presence of an	
CC	CC	epitope of Rhbg or RhBG in the sample. The nucleic acids of the invention	
CC	CC	are useful as probes for detecting transporter genes and particularly Rh	
CC	CC	type B transporter genes including e.g. NH 4 ⁺ ion transporters. The	
CC	CC	present nucleic acid sequence is that of the human RhBG gene located on	

CC chromosome 1q21.3. This sequence encodes the human RhBG protein of the
CC invention. RhBG is a polytypic transporter-type protein

XX Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 U; 0 Other;

Query Match 14.1%; Score 177; DB 6; Length 1805;
Best Local Similarity 49.5%; Pred. No. 3.8e-40;

Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCCAAGATCTGACCGGATGAGGGGGGCTTGGTGGCTTCTTCTACCTC 200
DB 218 CTACCCCAAGCTTCAGACGCTGCATGCTGCTTCTTGGGCTTGGGCTTCTTCTATGT 277
QY 201 GAATTTCCGAGACACAGCTGAGCAGTGTGGCTTCAACCTTTTCATGCTGGCGCTTGG 260
DB 278 CTTCGACAGCGTTTACGGCTTTCAGCAGCGTGGCTTCACTTCTCTCGCCGCTTTC 337
QY 261 TGTGACGGGCAATCTGCTGAGCGCTTCTTCAAGCCATTCCTCTTGGAGAGTGT 320
DB 338 CCGCAGTGTGTCACCTGCTCAGGGCTTCTTCACTCTTCTTCAAGGGTGGCCATCCCA 397
QY 321 CATCACTGTTCAATTCGCTGCGCCACATGAGTGTCTTGTGGTGTGATCTCAGT 380
DB 398 TGTGGCGTGGAGAGCATGATCAATGCTGACTTTGTGGGGGGCGCTGCTCATCTCTT 457
QY 381 GGATGCTGCTTGGGGAGAGTCACTTGGCGCATGTTGTGTGATGATGCTGTGAGAGT 440
DB 458 TGTGCGCTGCTGGGCAACCGGGCTAACCCAGCTGCTGCTCATGCGCTCTGAGAGT 517
QY 441 GACAGCTTTAGGCAACTGAGAGTGTATCATGTAATATCTTCAACAGACTACCAT 500
DB 518 GGTGCTGTTTGGCATCANTGAGTTGTGTGCTCTTATCTCTGGGGGTGAGAGATGCCG 577
QY 501 GAAKATGATGACATCTACGTGTGTGCAAGCTTATTTTGGGCTGTCTGTGGCTGTGCT 560
DB 578 AGGCTCCATGACTATCCACACTTTGTGCTTACTGCGGCTCGCTTGTGGGGTCT 637
QY 561 GCCAAGCCTTACCCGAGGGAGAGAGATTAAGATCAGACACAGATACCAGTTT 620
DB 638 GTACAGGCTCCAGCTGAGAGAGAGACAGCAGCGCTCCTCTACCATTCAGACT 697
QY 621 GTCTGCCATGCTGGCGGCTCTTCTTGTGTGATGTTCTGAGCAAGTTTCAACTGTGCT 680
DB 698 CTTCGCCATGATGGGACATCTTCTGTGTGATCTTCTGGCTTACCTTCAATGCTGACT 757
QY 681 GCTGAGAGTCCATTCAGAAAGAGATGCTGTCTTCAACACTTATGCTGTAGCAAT 740
DB 758 CACAGCGCTGGGGCTGGGAGCATGCGACGCGCTCCACACATATCTCCTGGCTGC 817
QY 741 CAGCGTGTGACACGCATCTCAGGGTATCTTGGCTCACCCCAAGGGAGATCAGCAA 800
DB 818 CAGCAGCTTGGACCTTTGCTTGTGTGAGCCCTTTAGGGAGAGATGGAGGCTTACAT 877
QY 801 GACTTATGTGACAGTCCGCTGTGTGAGAGCGCTGTGCTGTGGATACCTCTGTGCACT 860
DB 878 GTTCCATCATCAAAATGACAGCGCTGTGAGGGGTGTGTGGGACCTCAAGTAAAT 937
QY 861 GATCCCTTCTCCGATGCTGCATGAGTGTGCTGTTGGCTGGGCTGATCTCCGTCGG 920
DB 938 GATGCTGACACCTTTGGGGCTCTGGCAGCTGAGCTTCTTGGCTGGGACGTCTCCACGCT 997
QY 921 GGGAGCAGATACCTGCGGGGCTGTGTGAACGAGTGTGGGATTCGCCACAGCTCAT 980
DB 998 GGGGTACAAATTTTACGCGCCATCTCTTAATCAAAATTCAAAGTCAAGACATGTGG 1057
QY 981 CATGGCTACCAATTCAGCTTGTGGTGTGCTGTGAGAGATCATCTATGTTGCTGCT 1040
DB 1058 AGTCCCAACATCTCATGAGATGCGGGGGTCTTGGGGGCTCTCTGAGGGTCTTGTGCG 1117
QY 1041 GGTGCTTGAATACCTGTGAGC 1061
DB 1118 TGGACTTGGCACCACATGAAGC 1138

```

RESULT 10

AB211480 standard; cDNA; 1792 BP.

AC AB211480;

DT 20-JAN-2003 (first entry)

Human polynucleotide SEQ ID NO 362.

Human: genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiParkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; immunosuppressive; dermatological;
antiarthritis; gene; ss.

OS Homo sapiens.

PN WO200270539-A2.

PD 12-SEP-2002.

PF 05-MAR-2002; 2002WO-US005095.

PR 05-MAR-2001; 2001US-00799451.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

PI Weinman T, Wang J, Wang D, Drmanac RT;

DR WPI: 2002-759812/82.

DR P-PSDB; ABP69263.

PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.

PS Claim 1; SEQ ID NO 362; 1012bp + Sequence Listing; English.

CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB211119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, wound, burn, infection, ulcers, liver
CC platelet or coagulation disorders, myeloid or lymphoid disorders,
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIP0 at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;

Query Match 14.0%; Score 175.4; DB 6; Length 1792;
Best Local Similarity 49.4%; Pred. No. 1.1e-39;

Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCCAAGATCTGACCGGATGAGGGGGCTTGGTGGCTTCTTCTACCTC 200
DB 211 CTACCCCAAGCTTCAGACGCTGCATGCTGCTTCTTGGGCTTGGGCTTCTTCTATGT 270

```

```

QY 201 GAGTTTCCGAGACACAGCTGGAGCAAGTGGCTTCACTCTTCACTGCTGGCTTGG 260
DB 271 CTTCCTCAGAGCTTACCGCTTTCAGCAGCGTGGGCTTCACTCTCCCTGGCCGCTTTC 330
QY 261 TGTGAGTGGGCAATCCGTGCTGAGCGGCTTCCAGCCAGTTCCTTCTGGGAAGTGGT 320
DB 331 CTGCAAGTGGTCCACACTGGTCCAGGGCTTCTCCACTCTCTTCCACGGTGGCCATCCA 390
QY 321 CATCACACTGTTAGTATTCGGGCTGGCCACCATAGTCTTGTGCGGTGATCTCAGT 380
DB 391 TGTGGGCTGGAGAGCATGATCATGCTGACTTTGTGCGGGGCGCTGCTCATCTCTT 450
QY 381 GATGCTGTCTTGGGGAAGCTCAACTGGCGCAGTGGTGTGATGCTGCTGGAGGT 440
DB 451 TGGTGGCGTCTGGGCAAGACCGGGCTTACCACCTGCTGCTCATGCGCCCTGCTGAGGT 510
QY 441 GACAGCTTTAGGCAACTGAGGATGCTCATGAGTAATATTCTTCAACACAGCTACCAT 500
DB 511 GGTGCTGTGGCAATGATGATGCTTCTTCACTCTCTGGGGGTGAGAGATGCCGG 570
QY 501 GAACATGATGACACATCTACGTTTCCGAGCCTATTTTGGGCTGTCTGCGCCTGGTGG 560
DB 571 AGGCTCATGACTATCCACACACTTTGGTGGCTACTTGGGGCTGCTCTTGGCGGGTCT 630
QY 561 GCCAAGCCTCTACCCGAGGGAACGAGATAAAGATCAACAGCAACGATACCAGTTT 620
DB 631 GTACAGGCCCCAGCTGGAGAAAGCAAGACCGCCAGGGCTCCGCTTACCATTCAGACT 690
QY 621 GTCTGCATGCTGGGCCCTCTCTTGTGAGATTTCTGGCAGATTTCATCTGCTCT 680
DB 691 CTTCGCAATATTTGGACACATCTTCTGTGATCTTCTGGCTTGAAGCTTCAATGCTGACT 750
QY 681 GCTGAGAGTCAATCGAAAGAAAGATGCGGTTCACACACTACTATGCTGAGCAGT 740
DB 751 CACAGCCTGGGGCTGGGAGCATCGGACGCGCCCTCAACACATCTCTCTGGCTGC 810
QY 741 CAGGCTGTGACACCATCTTCAAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCA 800
DB 811 CAGCACCTTGGCACCTTTGGCTTGTCAAGCCCTTGTAGGGGAAGATGGAGGCTTGAAT 870
QY 801 GACTTATGTGACAGTGGCTGTGGTGGCAGAGGCGTGGCTGTGCTGTCTCACT 860
DB 871 GTTCAATCTCAAAATCAGCGCTGCTGAGGGGTTGTGTGGGACCTCAAGTGAAT 930
QY 861 GATCCCTTCTCGGCTGGCTGCAATGATGCTGAGTCTTGGGCTGGGCTGATCCGCTGC 920
DB 931 GATGCTGACACCTTTGGGGCTGTGGCAGCTGGCTTCTTGGCTGGGACTGTCTCAGCT 990
QY 921 GGAAGCCAAATACCTGCGGGGGTGTGTAACCGAGTGTGGGATTCGCCACAGCTCAT 980
DB 991 GGGGTACAAAGTTCTTCAAGCCCATCTTGAATCAAAATTCAAAGTCCAAAGACATGGG 1050
QY 981 CATGGGCTACAATTCACGCTTGTGGCTGTGCTTGGAGAGATCATCTAATTTGCTGCT 1040
DB 1051 AGTCCAAACCTCCATGGAATGCGGGGGTCTGGGGGCGCTCTGGGGGCTCTTGTGGC 1110
QY 1041 GGTGCTGATACCGTGGAGC 1061
DB 1111 TGGACTTGGCCACCATGAGC 1131

```

```

RESULT 11
ID ADM43998
XX ADM43998 standard; cDNA; 1792 BP.
XX
XX ADM43998;
XX
XX 03-JUN-2004 (first entry)
XX
XX Novel human arginine-rich protein cDNA #362.
XX
XX 88: gene; human; arginine-rich protein; cancer; inflammation;
KW genetic disorder.

```

```

XX XX Homo sapiens.
OS US2004053250-A1.
PN 18-MAR-2004.
PD 21-NOV-2002; 2002US-00302172.
PF 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002MO-US005095.
PR 20-AUG-2002; 2002US-00225251.
PA (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
PI Tang YT, Xue A, Drmanac RT;
PI WPI; 2004-238579/22.
DR
XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
PT Disclosure; SEQ ID NO 362; 51bp; English.
XX
XX The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
XX
SQ Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;

```

```

Query Match 14.0%; Score 175.4; DB 12; Length 1792;
Best Local Similarity 49.4%; Pred. No. 1.1e-39;
Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

```

```

QY 141 CTATCAAGTTGGCCAAATGATGACCGGTATGCGGCACTTGGGCTTCTCACTC 200
DB 211 CTACCCAGCTTCCAGAGCTGATGCCATGCTTGTGGCTTGAATCTTCTCATGCT 270
QY 201 GAGTTTCCGAGACACAGCTGAGCAGTGGCTTCAACCTTTCATGCTGGGCTTGG 260
DB 271 CTTCCTGACAGCTTACGGCTTTCAGACAGCTGGGCTTCACTTCTCTGGCGGCTTTC 330
QY 261 TGTGCAATGGGCAATCTGCTGAGCGGCTTCTGAGCAGTTCCTTCTGGGAAGTGGT 320
DB 331 CTGCAAGTGGTCCACACTGGTGTCCAGGGCTTCTCACTCTTCCACGAGTGGCCATCCA 390
QY 321 CATCACACTGTTAGTATTCGGGCTGGCCACCATAGTGTGCGGTGATCTCAGT 380
DB 391 TGTGGGCTGGAGAGCATGATCATGCTGACTTTGTGCGGGGCGCTGCTCATCTCTT 450
QY 381 GATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTGGTGTGATGCTGCTGAGGT 440
DB 451 TGGTGGCGTCTGGGCAAGACCGGGCTTACCACAGCTGCTGATGCGCCCTGAGAGT 510
QY 441 GACAGCTTTAGGCAACTGAGAGTGTGATCAGTAATATTCTTCAACACAGACTACCAT 500
DB 511 GGTGCTGTGGCAATGATGATGCTTCTTCACTCTCTGGGGGTGAGAGATGCCGG 570
QY 501 GAACATGATGACACATCTACGTTTCCGAGCCTATTTTGGGCTGTGCTGGCCTGGTGG 560
DB 571 AGGCTCATGACTATCCACACCTTTGGTGGCTACTTGGGCTGCTCTTTCGGGGTCT 630
QY 561 GCCAAGCCTCTACCCGAGGGAACGAGATAAAGATCAACAGCAACGATACCAGTTT 620

```

Db 631 GTACAGAGCCAGCTGAGAGAGCAAGCAGCGAGGCTCCGTACATTACAGCT 690
Qy 621 GTCTGCACTGCTGGAGCGCCCTCTTCTTGAGAGTCTGGCCAGTTTCACTGCTCT 680
Db 691 CTTCGCATGATTTGGAGCAATCTTCTTGAGATCTTGGCTTAAGCTTCAAGCTGCACT 750
Qy 681 GCTGAGAAAGTCCAAATGAAAGAAATGCGCTTCAACACCTACTATGCTGAGCACT 740
Db 751 CACAGCGCTGGGGGCTGGGAGCATGAGCGGCCCTCAACATACTCTCCCTGGCTGC 810
Qy 741 CAGCGTGTGACAGCATCTCAGGCTCATCTTGGCTCACCCCAAGGAGATGACAA 800
Db 811 CAGACCCCTTGGAGCATCTTGGCTCTTGTACGCCCTTGAAGGAGAAATGGAGGCTTGACAT 870
Qy 801 GACTTATGTGACAGATGCGGCTGTGAGAGAGCGCTGTGGGACTGCTGTCACCT 860
Db 871 GGTTCACATCTCAAAATGACAGCGCTGCTGAGAGGGGTGTGGTGGGACCTCAAGTGAAT 930
Qy 861 GATCCCTTCTCGTGGCTTGCATGCTGCTGGCTTGTGCTGGGCTGATCTCCGTCGG 920
Db 931 GATGCTGACACCCCTTGGGGCTCTGCGACCTGCTTCTTGGCTGGGACCTGCTCCAGCT 990
Qy 921 GGGAGCCAGTACTCTGCGGGGTGTGTAAACGAGGCTGGGAGTTCCCAAGCTCAT 980
Db 991 GGGGTACAAAGTCTTACAGCCCATCTTGAATCAAAATTCAAAGTCAAGACATGTGG 1050
Qy 981 CATGGCTTCAACTTCACTGCTGCTGGCTGTGCTTGGAGAGATCATCTTATGCTGCT 1040
Db 1051 AGTTCACAACTTCCATGAGATGCGGGGGTCTGAGGGCCCTCTCGGGGGTCTTGTGGC 1110
Qy 1041 GGTGCTGTATACGCTGGAGC 1061
Db 1111 TGGACTTGCACCCATGAAGC 1131
RESULT 12
ACD20392/C
ID ACD20392 standard; DNA; 1351 BP.
XX
XX ACD20392;
AC
XX
XX 26-AUG-2003 (first entry)
DT
XX
XX DNA encoding human NOV15 protein.
DE
XX
XX Human; NOVX; inflammatory disorder; demyelination disease; stroke;
KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;
KW hyperextension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;
KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;
KW haemophilia; autoimmune disease; allergy; AIDS;
KW graft versus host disease; Alzheimer's disease; arthritis; pain;
KW Parkinson's disease; Huntington's disease; obesity; diabetes;
KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;
KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;
KW noctropin; neuroprotective; cytosolic; antibacterial; virocidic;
KW protoscoleritis; antiatherosclerotic; hypotensive; cerebroprotective;
KW antiinflammatory; gynaecological; antiproliferative; dermatological;
KW hepatocarcinoma; haemostatic; immunosuppressive; antiallergic;
KW antiarthritic; anticomuleant; antiseborrhoeic; antiasthmatic;
KW neuroleptic; anti-HIV; analgesic; nephrotropic; antipsoriatic; Gene; ds.
XX
OS Homo sapiens.
XX
XX WO200298917-A2.
PN
XX
XX 12-DEC-2002.
PD
XX
XX 12-FEB-2002; 2002WO-US022049.
PF
XX
XX 12-FEB-2001; 2001US-0268221P.
PR 12-FEB-2001; 2001US-0268221P.
PR 13-FEB-2001; 2001US-0268466P.
PR 14-FEB-2001; 2001US-0268466P.
PR 14-FEB-2001; 2001US-0268466P.

PR 15-FEB-2001; 2001US-0269136P.
PR 16-FEB-2001; 2001US-0269310P.
PR 16-FEB-2001; 2001US-0269310P.
PR 15-MAR-2001; 2001US-0276405P.
PR 16-MAR-2001; 2001US-0276399P.
PR 16-MAR-2001; 2001US-0276703P.
PR 23-MAR-2001; 2001US-0278199P.
PR 28-MAR-2001; 2001US-0279274P.
PR 30-MAR-2001; 2001US-0280238P.
PR 02-APR-2001; 2001US-0280899P.
PR 08-AUG-2001; 2001US-0310797P.
PR 14-AUG-2001; 2001US-0312284P.
PR 14-SEP-2001; 2001US-0322284P.
PR 14-SEP-2001; 2001US-0322284P.
PR 18-OCT-2001; 2001US-0330293P.
PR 31-OCT-2001; 2001US-0335104P.
PR 31-OCT-2001; 2001US-0335109P.
PR 21-NOV-2001; 2001US-0332127P.
PR 28-NOV-2001; 2001US-0331772P.

(CURA-) CURAGEN CORP.

Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Lette M, Spytek KA;
Ji W, Casman SJ, Boldog FL, Patrajan M, Vernet CM, Ballinger RA,
Malayankar UM, Tchermey VT, Blalock AD, Gusev VY, Raschall L;
Mezes PD, Ellerman K, Heyes M, Herrmann JL, Shinkels RA, Iolme N;
Pena CE, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;
WPI; 2003-148650/14.
DR P-FSDB; ABO15004.
XX

Novel NOVX polypeptide useful for identifying an agent that binds to the
PT polypeptide and for treating cardiomyopathy, atherosclerosis,
PT hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel
PT disease.

Claim 9; Page 115; 566pp; English.

XX
XX The present invention relates to the isolation of novel human
XX polypeptides referred to as NOVX (NOV1-NOV37), variants of these
XX proteins, and the polynucleotide sequences encoding them. The NOVX
XX proteins of the invention share homology to various types of protein
XX families such as zinc finger-like proteins, enzymes, receptors, and
XX lipoproteins. The sequences of the invention may be useful in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease. For example, they can be used to treat inflammatory
XX disorders, demyelination disease, renal disorders, infections,
XX cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von
XX Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,
XX inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune
XX diseases, allergies, graft versus host disease, Alzheimer's disease,
XX arthritis, Parkinson's disease, Huntington's disease, obesity, diabetes,
XX acne, hair growth/loss, asthma, schizophrenia, AIDS, pain,
XX glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427
XX represent DNA sequences encoding the NOVX polypeptides of the invention.
XX Note: SEQ ID Nos 113-460 are known sequences used for homology purposes
XX

Sequence 1351 BP; 325 A; 372 C; 389 G; 265 T; 0 U; 0 Other;
Query Match 13.2%; Score 165.4; DB 8; Length 1351;
Best Local Similarity 49.1%; Pred. No. 7.5e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAGATGACCGGATGAGCGGCATTTGGGCTTCTCACTC 200
Db 1175 CTACCAAGCTTCCAGAGCGTACCGTATGCTTGTGGCTTCTCTCAATGAC 1116
Qy 201 GAGTTTCCGAGACACAGCTGAGCAAGTGTGGCTTCAACTTCTGATGCTGGCGTTGG 260
Db 1115 TTTCCTGACAGCGCTTACCGCTTTCAGCGCCGCTTCACTTCTGTTGGCAGCTTGG 1056
Qy 261 TGTGCACTGGGCAATCTGCTGAGCGCTTCTGAGCGCACTTCTTCTTGGAGAGTGGT 320

Db 1055 CATCCAGTGGGCGCTGCTCATGACGGGCTGGTTCACCTCTTACAGACCGCTACATCGT 996
 Qy 321 CATCACAGTGTTCAGTATTCGGCTGCGCACCATGAGTGTTCGCTGGTGAATTCAGT 380
 Db 995 CGTGGGGGTGGAGAACTCATCAACGCTGACTTCTGTGTGGCCCTCTCTCTGCGGTGCTT 936
 Qy 381 GGAATGCTGCTTGGGAGAGGTCAACTTGGCGCAGTGTGGTGTGATGCTGTGGAGGT 440
 Db 935 TGGGGGAGGTCTGGGTAAAGTCAAGCCCATTCAGCTGCTCATCATGACTTCTTCCAAGT 876
 Qy 441 GACAGCTTTAGGCACTGAGAGATGTCATCAGTAATATCTTCACACAGACTACCAT 500
 Db 875 GACCTCTTGTGCTGTGAATGATTCATCTTAACTCTGTAAAGGTGAAGGTGACAG 816
 Qy 501 GAACATGATGACATCTACGTTGTCGAGCTATTTTGGGCTGTGCTGTGCGCTGCTGCT 560
 Db 815 AGGCTCATATACATCCACACATTTGGCGCTACTTTGGGCTCAGAGTACCGGATCTT 756
 Qy 561 GCCAAGCCTCTACCCGAGGGAACGAGATAAAGATCAGACAGCAAGTACCAGTTT 620
 Db 755 CTACCGACGCACTAGACGAGACAGAGACAGACAATTTGTGTACAGTGGACCT 696
 Qy 621 GTCTGCATGCTGGGCGCCCTCTTCTGTGTGATGTTTGGCCAAAGTTTCAACTGCTCT 680
 Db 695 CTTTGCATATGATGGACCCCTCTTCTGTGTGATGTTTGGCCAAAGTTTCAACTGCTCT 636
 Qy 681 GCTGAGAGTCAATCGAAAGAAAGATGCGGTGTTCAACACTACTATGCTGTGAGAGT 740
 Db 635 ATCTTACCATGGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 576
 Qy 741 CACGCTGTGACAGCCATCTCAGGGTCACTTGTGGCTCACCCCAAGGAGATCAGCAA 800
 Db 575 CTGCGTCTTACCTCGGTGGCAATATCAAGTCCCTGCAAGAAAGGCAAGCTGACAT 516
 Qy 801 GACTTATGTGACAGTCCGCTGTGTGACAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 860
 Db 515 GGTGCAATCAGAAATCCAGCTGCGAGAGAGGGGTGGCGGTGGTGTGTGTGTGTGTGT 456
 Qy 861 GATCCCTTTCGCGGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 920
 Db 455 GATGCTATGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 Qy 921 GGGAGCCAGTACCTGCGGGGGGTGTGTAAACGAGTCTGGGGATTTCCCAAGCTTCAT 980
 Db 395 GGGTTTGTATACCTGACCCCATTCCTGAGAGTCCCGGCTGACATTCAGAGACATGTGG 336
 Qy 981 CATGGCTACAACTTCACTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1035
 Db 335 CATTAACAAATCTGCATGCGCATTTCTGCGCATCATAGGCGGATCTGTGGTGTGTGTG 281

RESULT 13

AAC64286
 ID AAC64286 standard; cDNA; 1853 BP.

AAC64286;

23-FEB-2001 (first entry)

Human membrane-associated protein HUMAP-13 cDNA.

Human membrane-associated protein; HUMAP; transgenic organism;
 drug screening; cell signalling modulator; agonist; antagonist;
 cell differentiation modulator; cell proliferation modulator;
 cell proliferation disorder; cancer; cell differentiation disorder;
 developmental disorder; cell signalling disorders; endocrine disorder;
 hyperparathyroidism; hyperparathyroidism; infection;
 pancreatic disorder; diabetes mellitus; immunological disorder;
 hereditary neuropathy; gonadal steroid hormone associated disorder;
 infertility; ss.

Homo sapiens.

PN WO200065054-A2.
 XX 02-NOV-2000.
 PD 20-APR-2000; 2000MO-US010884.
 XX 23-APR-1999; 99US-0130694P.
 PR 23-JUN-1999; 99US-0140580P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Hillman JL, Bandman O, Tang YT, Lai P, Yue H, Reddy R;
 PI Azimzai Y, Baughn MR;
 PI WPI; 2000-687346/67.
 DR P-PSDB; AAB29656.
 XX Human membrane-associated protein, useful for diagnosis and treatment of
 PT cell signalling, cell differentiation and cell proliferation disorders
 PT such as cancer, and for identifying agonists and antagonists.
 PS Claim 4; Page 96; 99p; English.

The invention relates to 17 human membrane-associated proteins, HUMAP-1
 to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them (AAC64274-
 C64290). The invention also relates to expression constructs, host cells
 and transgenic organisms comprising a HUMAP nucleic acid sequence; the
 recombinant preparation of a HUMAP; methods of screening compounds for
 their ability to modulate HUMAP activity or expression; and
 pharmaceutical compositions comprising a HUMAP protein, a HUMAP agonist
 or HUMAP antagonist. The HUMAPs acts as modulators of cell signalling,
 differentiation and proliferation. A HUMAP is useful for screening a
 compound for effectiveness as an agonist or antagonist of HUMAP activity.
 The protein, or the identified agonist or antagonist is useful for
 treating a disease or condition associated with decreased or increased
 expression of functional HUMAP. A HUMAP nucleic acid is useful for
 screening a compound for its ability to alter expression of that
 particular HUMAP gene. A wide variety of disease may be treated using
 compositions of the invention. These diseases include cell proliferative
 disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g.,
 breast, bladder, bone marrow, brain and uterine cancer); cell
 differentiation disorders, in particular developmental disorders (e.g.,
 renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,
 epilepsy, and muscular dystrophy); cell signalling disorders, in
 particular endocrine disorders such as hypohalms and pituitary
 disorders resulting from lesions such as thrombosis, disorders associated
 with hyperparathyroidism (e.g., acromegaly); disorders associated with
 hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders
 such as type I or type II diabetes mellitus; infections; immunological
 disorders; hereditary neuropathies (e.g., neurofibromatosis); and
 disorders associated with gonadal steroid hormones (e.g., infertility,
 endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell
 deficiency and gynecomastia). Antibodies which specifically bind HUMAP
 may be used for the diagnosis of disorders associated with the expression
 of HUMAP, or in assays to monitor patients being treated with HUMAP or
 agonists, antagonists or inhibitors of HUMAP. The present sequence
 CC represents a HUMAP cDNA of the invention

Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 U; 0 Other;

Query Match 13.2%; Score 165.4; DB 3; Length 1853;
 Best Local Similarity 49.1%; Pred. No. 8.9e-37;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAAGATTTGACCGTATGCGGCCCATTTGGGCTTCTTCACCTC 200
 Db 207 CTACCCAGGCTTCAGAGACCTGACGCTGAGTCTTCTGTGGGCTTCTTCATGAC 266
 Qy 201 GAATTTCCGGAGACAGAGCTGAGCAGTGGGCTTCAACCTTCAATGCTGGGCTTGG 260
 Db 267 TTTCTGCAAGCTTACGCTTACGCGCGGCTTCAACTTCTCTGTGGCAGGCTTGG 326
 Qy 261 TGTGAGTGGGCAATCCTGCTGAGCGGCTTCTGAGCCAGTTCCTTGTGGAAAGTGTGT 320

```

Db      327 CATCAGTGGCGCTGCTCATGACAGGCTGTGTCACCTTCTTAACAAGACCGCTACATCGT 386
Qy      321 CATCACTGTGTGATGATTCGGCTGGCAACATGAGTGTCTTGGTCTGATCTCACT 380
Db      387 CGTGGCGGTGAGAACTCATCAACGCTGATCTTGGCGCTCTGTGTGGTGGCTCT 446
Qy      381 GATGCTGTCTTGGGAAAGTCAACTTGGCGAGTGGTGTGATGATGATGATGATGATGAT 440
Db      447 TGGGGGAGTTCTGGTAAAGTACGCCCATTTAGCGTCTCATCATGACTTTCTTCAAGT 506
Qy      441 GACAGCTTAAAGCACTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 500
Db      507 GACCTCTTGGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
Qy      501 GAAATGATGACATCTAGTGTGACAGCTATTTTGGGCTGTGCTGGCTGGCTGGCTGG 560
Db      567 AGGCTCATGACCATCAACATTTGGCGCTACTTTGGGCTCAAGATGACCGGATCTT 626
Qy      561 GCCAAAGCTTACCCGAGGGAACGAGAGATTAAGATCAAGCAAGATACCACTTT 620
Db      627 CTACGACGCAACTAGACAGAGCAAGAGAGACAGAACTTGTGTACAGTGGACCT 686
Qy      621 GTCTGCATGCTGGGCGCTCTTCTTGTGATGATGATGATGATGATGATGATGATGAT 680
Db      687 CTTCGCATGATGGCACTTCTTCTGTGATGATGATGATGATGATGATGATGATGATGAT 746
Qy      681 GCTGAAGATCCATTCGAAAGAGAGATGCGGTGTTCACACTATGCTGTAGCAGT 740
Db      747 ATCTTACCATGGGACAGCCAGCACGAGCGCCCATCAACACTACTCTCTTGGAC 806
Qy      741 CAGCGTGGTGAAGCATCTCAAGGTCTACTTGGCTCAACCCCAAGGAAATCAGCA 800
Db      807 CTGGGTCTTACTTGTGGCAATATCCAGTCCCTGCAAAAGGCAAGCTGAGCAT 866
Qy      801 GACTTATGTGCAAGTGGGTGTGTGGCAGAGGCGTGTGGTGTGCTGTGCTGTGCT 860
Db      867 GGTGACATTCAGATATGCAAGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 926
Qy      861 GATCCCTTCTCGGTGGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 920
Db      927 GATGCTATGCTTAAAGTGGTCCCTCATCATGCTGCTGTGCTGCGGATCATCTCAACCT 986
Qy      921 GGGAGCCAAATGCTGCGGGGTGTGTAACCGAGTCTGGGATTTCCCAAGCTTCAT 980
Db      987 GGGTTTGTATCTGACCCATCTCTGAGTCCCGGCTGCAATCCAGGACATGTGG 1046
Qy      981 CATGGGCTAACTTCAAGCTTGTGGGTCTGCTTGAAGATCATCTAATGTG 1035
Db      1047 CATTAAGATCTGCATGGCATTCCTGGCATCATAGGCGGATGTGGGTGTGTG 1101

```

RESULT 14
AAH25763
ID AAH25763 standard; cDNA; 1949 BP.

XX AAH25763;

DT 14-AUG-2001 (first entry)

DE Oesophagus cancer associated DRC2 coding sequence.

KW DRC2; oesophagus cancer; ss.

XX Unidentified.

XX Key Location/Qualifiers

FT CDS 41..1480 /tag= a

FT /product= "DRC2"

XX CN1283694-A.

```

PD      14-FEB-2001.
XX
XX      10-AUG-1999; 99CN-00117523.
PF
XX      10-AUG-1999; 99CN-00117523.
PR
XX      (ONCO-) ONCOLOGY INST TUMOR HOSPITAL CHINESE ACA.
PA
XX      Wang M, Xu Z, Xu X;
XX      WPI; 2001-291757/31.
DR      P-PSDB; AAB97000.
XX
XX      Gene associated with esophagus cancer.
PT
XX      Claim 4; Fig 1; 33pp; Chinese.
XX
XX      The present invention provides the protein and coding sequences of DRC2,
CC      which is associated with esophagus cancer. Also provided is a method for
CC      diagnosing diseases associated with an abnormal version of the nucleic
CC      acid and encoded protein, involving detecting any nucleic acid sequence
CC      mutations, methylation and variation at the RNA and polypeptide level.
CC      The sequences can be used in the treatment of cancer.
SQ
Sequence 1949 BP; 382 A; 593 C; 530 G; 444 T; 0 U; 0 Other;

```

Query Match 13.2%; Score 165.4; DB 5; Length 1949;
Best Local Similarity 49.1%; Pred. No. 9,1e-37;
Matches 499; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```

Qy      141 CTATCAAGTTGGCCCAAGATCTGACCGTATGGCGGCAATGGCTTGGCTTCTCACTC 200
Db      217 CTACCAAGCTTCCAGAGAGTGCATGATGATCTTCTGGGCTTGGCTTCTCACTAC 276
Qy      201 GAGTTTCCGAGAACACAGCTGAGAGATGTGTGCTTCAACTCTTCACTGTGGGCTTGG 260
Db      277 TTTCCTGACGCGCTACGCGCTTACGCGCGGTGCTTCAACTTCTGTTGGAGCTTGG 336
Qy      261 TGTGAGTGGGCAATCTGCTGAGACGCGTCTGAGCCAGTTCCTTGTGAAAGTGT 320
Db      337 CATCACTGGGCGCTGCTCATGACAGGCTGTGCTTCACTTCAAGACCGCTACATCT 396
Qy      321 CATCACTGTTCAATATTCGCTGGCCACCATGATGATCTTGTGGGTGATCTCAGT 380
Db      397 CGTGGCGGTGAGAACTCATCAACGCTGATCTTGGGTGCTTGTGCGGTCTT 456
Qy      381 GGAATGCTCTTGGGAAAGTCACTTGGCCAGTGTGTGTGATGATGATGATGATGATGAT 440
Db      457 TGGGCACTTCTGGGTAAAGTCAACCCCATTCACCTCTCATCAATGACTTCTTCCAAGT 516
Qy      441 GACAGCTTAAAGCAACTGAGATGATCATCAGTATATCTTCAACACAGACTACCAT 500
Db      517 GACCCCTTGGCTGTGATGATGATCTTCTTAACTGTAAAGTGAAGATGACAG 576
Qy      501 GAAATGATGACATCTAAGTGTGCAAGCTATTTTGGCTGTGTGGCTGTGGCTGTGGCT 560
Db      577 AGGCTCATGACCATCAACATTTGGCGCTTACTTTGGGCTCAAGTGAACCGGATCT 636
Qy      561 GCCAAAGCTTACCCGAGGGAACGAGAGATTAAGATCAACAGCAAGATACCACTTT 620
Db      637 CTACCAAGCAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
Qy      621 GTCTGCATGCTGGGCGCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 680
Db      697 CTTCGCATGATGGCACTTCTTCTGTGATGATGATGATGATGATGATGATGATGATGAT 756
Qy      681 GCTGAAGATCCATTCGAAAGAGAGATGCGGTGTCAACACTTATGCTGTACAGT 740
Db      757 ATCTTACCATGGGACAGCAGCAGAGCGCCCATCAACACTTATGCTTGTGGCAGC 816
Qy      741 CAGCGTGGTGAAGCATCTCAAGGTATCTTGTGCTCAACCCCAAGGAAAGTATGACAA 800
Db      817 CTGGGTCTTACTCGGTGGCAATATCAAGTCCCTGCAAGAGAGGCAAGCTGAGCAT 876

```


QY 801 GACTTATGTGCACAGTGCCTGTTGGAGAGGCGTGTGGTACCTTCGTCACT 860
DB 877 GGTGCACATCCAGAAATGCCAGCTCCGAGAGGGGTGCGCGTGGTACCGCTGTGAGAT 936
QY 861 GATCCCTTCCCGGCGTTCGCATGTGCTGGGTCTTGTGGCTGTGATCTCCGTGG 920
DB 937 GATGCTCATGCTTACCGTCCCTCATCATCGGCTTGTGCTGGCATCATCTCCACCT 996
QY 921 GGGAGCCAGTACCTGCGGGGTGTGTAAACCGAGTCTGGGAGATTCCCCACAGCTCCAT 980
DB 997 GGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGGCTGCACATCCAGACATGTGG 1056
QY 981 CATGGGCTACAACTTCACTGCTGTGGGTCTGCTTGAAGATCATCTTCAATTGTG 1035
DB 1057 CATTAACATCTGCATGTGCATCTCGGCATCATGAGGGGAGATGGTGGTGTGTG 1111

RESULT 15
ABK49215
ID ABK49215 standard; cDNA; 1952 BP.
XX
AC ABK49215;
XX
DT 15-JUL-2002 (first entry)
XX
DE cDNA encoding human Rh type C gene (RHCG) protein.
XX
KW RHCG: human; non-erythroid Rh type C glycoprotein; chromosome 15q25;
XX gene; 88.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 24..1440
FT /tag= a
FT /product= "RHCG protein"
FT polyA_signal 1885..1890
FT /tag= b
XX
PN M0200220719-A2.
XX
PD 14-MAR-2002.
XX
PP 05-SEP-2001; 2001MO-US027503.
XX
PR 07-SEP-2000; 2000US-0230660P.
XX
PA (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI Huang C, Liu Z;
XX
XX WPI; 2002-351774/38.
XX P-PSDB; AAU78997.
XX
PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C
PT and glycoproteins which have a characteristic twelve transmembrane domain
PT structure.
XX
XX Claim 2; Fig 1; 53pp; English.
XX
XX This invention relates to the nucleic acid and protein sequences of novel
XX human and mouse non-erythroid Rh type C glycoprotein (RHCG). The RHCG
XX protein and the mouse homologue (rhcg) have a characteristic 12
XX transmembrane domain structure and are expressed in kidneys and testis.
XX The invention also comprises a method for antibody that specifically
XX binds an epitope of the glycoprotein and a method for detecting the
XX protein using this antibody. The antibodies of the invention may be used
XX in Western blot, enzyme linked immunosorbent assays (ELISA) or
XX immunohistochemical assays to identify the non- erythroid tissues,
XX particularly kidney and testis, that express the RHCG or Rhcg
XX glycoproteins. The methods are used for detecting an Rhcg or RhCG
XX glycoprotein in a sample. The present sequence represents the cDNA

CC encoding the human Rhcg Rh type C glycoprotein (RhCG) protein sequence of
CC the invention. The gene encoding this protein is located on human
CC chromosome 15q25
XX
SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 U; 0 Other;
Query Match 13.2%; Score 165.4; DB 6; Length 1952;
Best Local Similarity 49.1%; Pred. No. 9.1e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
QY 141 CTATCAAGTTGGCCAGATATGACCTGATGAGGCGCATTTGGGCTTCTTCACTC 200
DB 201 CTACCCAGCTTCAGAGACGTGACCGTGAAGTCTTGTGGGCTTCCATATGAC 260
QY 201 GAGTTTCCGAGACACAGCTGAGAGAGTGGGCTTCAACTCTTCATAGCTGGGCTTGG 260
DB 261 TTTCCTGCAAGCTTACGAGCTTACGCGCCGCTTCAACTTCTGTTGGCAAGCTTGG 320
QY 261 TGTGCAATGGGCAATCTCTGTGACGAGCTTCTGAGCCAGTTCCTTGGGAAGTGTG 320
DB 321 CATCCAGTGGGCGCTGTCTATGACAGGCGTGTCCACTTTCAGAGCCGTAATCGT 380
QY 321 CATCAACTGTTCATATTTGGCTGGCCACATGAGTCTTGTGGGTGTATTCAGT 380
DB 381 CGTGGGCGTGGAGAACCTCATCAAGCTGACTTCTGTGGCCCTCTGTGGCGGCTT 440
QY 381 GATGCTGTCTTGGGGAGGTCAACTTGGGCGAGTTGGTGGTGGTGGTGGAGT 440
DB 441 TGGGGAGTTTCTGGTAAATGACGCCCATTCAGCTGTATATGACTTTCTTCCAGT 500
QY 441 GACAGCTTAAAGCACTGAGATGTGATCAGTAATATCTTCAACACAGACTACACAT 500
DB 501 GACCTCTTGGCTGTGAATGATTCATTCCTTAACTGTGAAGTGAAGATGACAG 560
QY 501 GAACATGATGACATATCTACGTGTTCCAGGCTTATTTGGGCTGTGGGCTGTGGCT 560
DB 561 AGGCTCATGACCATCCACACATTTGGCGCTCTTTGGGCTCAGATCCGGATCTT 620
QY 561 GCCAAGCCTCTACCCGAGGGAAGGAGTAAGATCAAGACAGCAAGATACCAAGTTT 620
DB 621 CTACCGAGCAACTTGAAGACAGACAGAGAGACAGAAATTTGTGACAGTGGACCT 680
QY 621 GTCTGCATGCTGGGCGCCCTCTTCTTGTGATGTTCTGGCAAGTTTCACTCTGCT 680
DB 681 CTTTGCATGATTTGGACCCCTCTTCTGTGAGATGTACTGGCCAGCTTCACTGACCAT 740
QY 681 GCTGAGAGTCCAAATGAAAGAGAAATGCGTGTTCACACTTATCTGTAGCAT 740
DB 741 ATCTACCATGGGAGACAGACAGACCGAGCCCATCAACCTTACTGTCTTGGGAGC 800
QY 741 CAGCGTGGAGACAGCATCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATACAGCA 800
DB 801 CTGCGTGTCTTACTCTGCTGCAATATTCAGTGCCTTGCACAGAGGCAAGCTGACAT 860
QY 801 GACTTATGTGCACAGTGCCTGTTGGACAGAGCGTGTGGTACCTCGTGTCACT 860
DB 861 GGTGCACATCCAGAAATGCCAGCTCCGAGAGGGGTGGCCGTGGTACCGCTGTGAGAT 920
QY 861 GATCCCTTCCCGTGTTCGCATGTGCTGGGTCTTGTGGCTGTGATCTCCGTGG 920
DB 921 GATGCTCATGCTTACCGTCCCTCATCATCGGCTTGTGTGGGATCATCTCCACCT 980
QY 921 GGGAGCCAGTACCTGCGGGGTGTGTAAACCGAGTCTGGGAGATTCCCCACAGCTCCAT 980
DB 981 GGGTTTGTATACCTAACCCCATTTCTGGAGTCCGCGCTGCACATCCAGACATGTGG 1040
QY 981 CATGGGCTACAACTTCACTGCTGTGGGTCTGCTTGAAGATCATCTTCAATTGTG 1035
DB 1041 CATTAACATCTGCATGTGCATCTCGGCATCATGAGGGGAGATGGTGGTGTGTG 1095

Search completed: March 23, 2005, 01:08:38
Job time : 709 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1252.4	99.9	1354	18	US-10-672-764A-35	Sequence 35, Appl1	
2	365.8	29.2	487	10	US-09-918-995-30694	Sequence 30694, Appl1	
3	175.4	14.0	1792	17	US-10-302-172-362	Sequence 362, Appl1	
c	4	165.4	13.2	1351	17	US-10-074-978A-41	Sequence 41, Appl1
5	165.4	13.2	1440	9	US-09-949-145-5	Sequence 5, Appl1	
6	165.4	13.2	1952	9	US-09-949-145-1	Sequence 1, Appl1	
7	165.4	13.2	1952	17	US-10-172-118-1635	Sequence 1635, Appl1	
8	165.4	13.2	1952	17	US-10-342-887-1635	Sequence 1635, Appl1	
9	150.2	12.0	1497	9	US-09-949-145-6	Sequence 6, Appl1	
10	150.2	12.0	2097	9	US-09-949-145-2	Sequence 2, Appl1	
c	11	148.8	11.9	2415	17	US-10-104-047-730	Sequence 730, Appl1

C	12	144.6	11.5	505	13	US-10-027-632-282039	Sequence 282039, Ap
C	13	144.6	11.5	505	13	US-10-027-632-282040	Sequence 282040, Ap
C	14	144.6	11.5	505	17	US-10-027-632-282039	Sequence 282039, Ap
C	15	144.6	11.5	505	17	US-10-027-632-282040	Sequence 282040, Ap
C	16	144.2	11.5	505	13	US-10-027-632-282041	Sequence 282041, Ap
C	17	144.2	11.5	505	13	US-10-027-632-282042	Sequence 282042, Ap
C	18	144.2	11.5	505	17	US-10-027-632-282041	Sequence 282041, Ap
C	19	144.2	11.5	505	17	US-10-027-632-282042	Sequence 282042, Ap
C	20	141.4	11.3	681	13	US-10-027-632-136609	Sequence 136609, Ap
C	21	141.4	11.3	681	17	US-10-027-632-136609	Sequence 136609, Ap
C	22	139.8	11.1	486	9	US-09-864-761-1547	Sequence 1547, Ap
C	23	120	9.6	1571	17	US-10-264-237-1007	Sequence 1007, Ap
C	24	111.8	8.9	123	9	US-09-864-761-18306	Sequence 18306, Ap
C	25	109.6	8.7	1650	14	US-10-012-542-75	Sequence 75, Ap
C	26	109.6	8.7	1650	14	US-10-011-123-75	Sequence 75, Ap
C	27	108.6	8.7	15918	15	US-10-911-455-1590	Sequence 1590, Ap
C	28	101.2	8.1	1765	17	US-10-190-115-21	Sequence 21, Ap
C	29	101.2	8.1	1765	17	US-10-369-072-21	Sequence 21, Ap
C	30	82.2	6.6	16918	15	US-10-311-455-1589	Sequence 1589, Ap
C	31	79.2	6.3	730	17	US-10-264-237-1194	Sequence 1194, Ap
C	32	63	5.9	489	10	US-09-918-595-1169	Sequence 1169, Ap
C	33	61.2	4.9	198	11	US-09-864-408A-1671	Sequence 1671, Ap
C	34	57	4.5	449	14	US-10-052-283-105	Sequence 105, Ap
C	35	56.6	4.5	446	9	US-09-864-761-2069	Sequence 2069, Ap
C	36	53.6	4.3	505	16	US-10-029-386-9596	Sequence 9596, Ap
C	37	53.4	4.3	198	16	US-10-029-386-23296	Sequence 23296, Ap
C	38	51	4.1	1493	16	US-10-029-386-25133	Sequence 25133, Ap
C	39	48.6	3.9	1973	9	US-09-864-761-2471	Sequence 3471, Ap
C	40	48.6	3.9	2543	16	US-10-029-386-20536	Sequence 20536, Ap
C	41	48	3.8	401	9	US-09-864-761-1936	Sequence 3936, Ap
C	42	48	3.8	806	13	US-10-027-632-158717	Sequence 158717, Ap
C	43	48	3.8	806	17	US-10-027-632-158717	Sequence 158717, Ap
C	44	47.8	3.8	1152	17	US-10-282-122A-18037	Sequence 14837, Ap
C	45	47.8	3.8	1972	15	US-10-017-161-1981	Sequence 1981, Ap

ALIGNMENTS

```

RESULT 1
US-10-672-764A-35
; Sequence 35, Application US/10672764A
; Publication No. US20040156832A1
; GENERAL INFORMATION:
; APPLICANT: Jolly, Chris
; TITLE OF INVENTION: Immunoglobulin Compositions and Methods
; FILE REFERENCE: 1331.1001U
; CURRENT APPLICATION NUMBER: US/10/672,764A
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1354
; TYPE: DNA
; ORGANISM: Human
US-10-672-764A-35

```

Query Match	99.9%;	Score 1252.4;	DB 18;	Length 1354;
-------------	--------	---------------	--------	--------------

```
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	1	ATAGAGCTCAAGTAAACCCGGGAGTCTGTCCGGGGGCTGCCTCTGTGGGCGCTTAACACTG	60
Db	1	ATAGAGCTCAAGTAAACCCGGGAGTCTGTCCGGGGGCTGCCTCTGTGGGCGCTTAACACTG	60
QY	61	GAGGAGGCTCATTTCTCCTCTTCAATTTTTTTTACCACTATGACGGTTCCTTAGAGAT	120
Db	61	GAGGAGGCTCATTTCTCCTCTTCAATTTTTTTTACCACTATGACGGTTCCTTAGAGAT	120
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAGATCTGACCGTGAATGGGGCCATT	180
Db	121	CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAGATCTGACCGTGAATGGGGCCATT	180

QY 181 GGGCTTGGGCTTCTCTCACTGAGTTTCCGGAGACAGAGCTGGAGAGAGTGGCTTCAAC 240
DB 181 GGGCTTGGGCTTCTCTCACTGAGTTTCCGGAGACAGAGCTGGAGAGAGTGGCTTCAAC 240
QY 241 CTTTCAATGCTGGCGCTTGGTGTGCAAGTGGCAATCTTCTGAGCGCTTCTTGAGCCAG 300
DB 241 CTTTCAATGCTGGCGCTTGGTGTGCAAGTGGCAATCTTCTGAGCGCTTCTTGAGCCAG 300
QY 301 TTCCCTTGGGAAAGGTGTGATCAACAGTTCAGATTCGGCTGGCCACCATGAGTGTCT 360
DB 301 TTCCCTTGGGAAAGGTGTGATCAACAGTTCAGATTCGGCTGGCCACCATGAGTGTCT 360
QY 361 TTGCTGGTCTGATCTCAAGTGTGATCTTGGGAAAGTGTCAATTGGCCAGTGTG 420
DB 361 TTGCTGGTCTGATCTCAAGTGTGATCTTGGGAAAGTGTCAATTGGCCAGTGTG 420
QY 421 GTGATGTGCTGGTGGAGGTGACAGTTTAAAGCACTGAGAGTGTATCATGATATTC 480
DB 421 GTGATGTGCTGGTGGAGGTGACAGTTTAAAGCACTGAGAGTGTATCATGATATTC 480
QY 481 TTCAAACAGACTACACATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTCAAACAGACTACACATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGTGGCTGTGGTGTGCAAGCTTCAACCGGAGGAGCGAGATGATGATGATG 600
DB 541 CTGTCTGTGGCTGTGGTGTGCAAGCTTCAACCGGAGGAGCGAGATGATGATGATG 600
QY 601 ACAGACAGATACCCAGTTTGTGCTGCAATGCTGGGCGCTTCTTGTGATGATGATG 660
DB 601 ACAGACAGATACCCAGTTTGTGCTGCAATGCTGGGCGCTTCTTGTGATGATGATG 660
QY 661 CCAAGTTTCAACTGTCTGTGCTGAGAAAGTCAATGCAAGAAAGAAAGTGGCTTCAAC 720
DB 661 CCAAGTTTCAACTGTCTGTGCTGAGAAAGTCAATGCAAGAAAGAAAGTGGCTTCAAC 720
QY 721 ACTACTATGCTGATGAGAGTCAAGGCTGTGACAGCCATCTCAGGCTATCTTGGCTCAC 780
DB 721 ACTACTATGCTGATGAGAGTCAAGGCTGTGACAGCCATCTCAGGCTATCTTGGCTCAC 780
QY 781 CCCCAGGAGAAAGATCAGCAAGACTTATGTGACAGTGGTGTGAGAGAGGCGTGGCT 840
DB 781 CCCCAGGAGAAAGATCAGCAAGACTTATGTGACAGTGGTGTGAGAGAGGCGTGGCT 840
QY 841 GTGGGTATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 GTGGGTATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 GCTGGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGAGGTGTGTAACGAGTGTG 960
DB 901 GCTGGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGAGGTGTGTAACGAGTGTG 960
QY 961 GGGATTTCCCAAGCTTCATCATGAGGCTCAACTTCACTTGTGCTGGAGAG 1020
DB 961 GGGATTTCCCAAGCTTCATCATGAGGCTCAACTTCACTTGTGCTGGAGAG 1020
QY 1021 ATCATCTACATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 ATCATCTACATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 TTCCAGGCTCTCTCTGAGATTGGGAACTCAGTTGGCCATGTATAGCTTCACTGCT 1140
DB 1081 TTCCAGGCTCTCTCTGAGATTGGGAACTCAGTTGGCCATGTATAGCTTCACTGCT 1140
QY 1141 GGTCTCTCAAGAGTTTGTCTCTTAATCTTAAATATGGAAGCACTTCATGAGGCTTAA 1200
DB 1141 GGTCTCTCAAGAGTTTGTCTCTTAATCTTAAATATGGAAGCACTTCATGAGGCTTAA 1200
QY 1201 TATTTTGAACCAAGTTTCTGAAAGTTTCTCTCACTTGGCTGTGGAATTTTAA 1254
DB 1201 TATTTTGAACCAAGTTTCTGAAAGTTTCTCTCACTTGGCTGTGGAATTTTAA 1254

RESULT 2
US-09-918-995-30694
; Sequence 30694, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: FROM VARIOUS CDNA LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30694
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(487)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30694

Query Match 29.2%; Score 365.8; DB 10; Length 487;
Best Local Similarity 99.5%; Pred. No. 2,4e-106;
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 706 AATGCCGATGTTCAACCACTACTATGCTGTAGCAGTCAAGTGGAGAGCCATCTCAGGG 765
DB 56 ATTCCCGTGTCAACACTTATGCTGTAGCAGTCAAGTGGAGAGCCATCTCAGGG 115
QY 766 TATCTTGGCTCACCCCAAGGAAATCAGCAAGCTTATGTGACAGTGGCTGTG 825
DB 116 TCATCTTGGCTCACCCCAAGGAAATCAGCAAGCTTATGTGACAGTGGCTGTG 175
QY 826 GCAGAGAGCGTGGCTGTGGTGAACCTCGTGTGACCTGATCCCTTTCGCTGGCTTGCATG 885
DB 176 GCAGAGAGCGTGGCTGTGGTGAACCTCGTGTGACCTGATCCCTTTCGCTGGCTTGCATG 235
QY 886 GTGCTGGGCTTGTGCTGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGAGGT 945
DB 236 GTGCTGGGCTTGTGCTGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGAGGT 295
QY 946 TGTAAACGAGTCTGGGAGTTTCCCAACAGCTCCATCATGAGGCTTCAACTTCACTTGTGCTG 1005
DB 296 TGTAAACGAGTCTGGGAGTTTCCCAACAGCTCCATCATGAGGCTTCAACTTCACTTGTGCTG 355
QY 1006 GGTCTGCTGGAGATCATCTACATTTGTGCTGTGGTGTGATACCGTGGAGCGGC 1065
DB 356 GGTCTGCTGGAGATCATCTACATTTGTGCTGTGGTGTGATACCGTGGAGCGGC 415
QY 1066 AATGGCATG 1074
DB 416 AATGGCATG 424

RESULT 3
US-10-302-172-362
; Sequence 362, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aiding J.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803, 1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20

Qy	141	CTATCAAGTTGGCCAAAGTCTGACCGTGATGGCGGCATTTGGCTTGGCTTCTCACTC	200
Db	211	CTACCCAAAGCTTCCAGAGCGTCATGCCATGGTCTTCCGTGGCTTTGACTTCTCATGGT	270
Qy	201	GAGTTTCCGAGACACACAGCTGGAGCAGTGGCCCTTCAACCTCTTCATGTCGGCCTTGG	260
Db	271	CTTCTCGAGGGTTAACGGCTTCAGAGAGGTGGGCTTCACTTTCCTGGCGCCCTTGGC	330
Qy	261	TGTGCAAGTGGGCAATCCTGCTGAGACGGCTTCTTGAGCCAGTTCCCTTCTGGAGAGTGT	320
Db	331	CTTCGAGTGGTCCACACTGCTTCAGGGGCTTTCTCCACTCTTCCACGGGTGGCCATACCA	390
Qy	321	CATCAACCTGTTCACTATTTGGCTGGGACCAATGATGTCCTTTGTGGGTGCTGATCTCACT	380
Db	391	TGTTGGCGGTGAGAGCATGATCATATGCTGACTTTTGTGCGGGGGCGGTGCTCATCTCTT	450
Qy	381	GGATGCTGTCTTGGGGAAAGTCAACTTTGGCGCAGTTGGTGGATGATGATGCTGGAGGT	440
Db	451	TGTGTCCGCTCTGGGCAAGACCGGGCCATCCAGCTGTGCTCAATGGCCCTGTCTGAAGT	510
Qy	441	GACAGCTTTAGGCAACTGAGAGTGTATCATGATAATCTTCAACACAGACTTACCAT	500
Db	511	GGTCTGTTTGGCATCAATGATTTGTCTCTCTTCATCTCTGGGGGTAGAGATGCCGG	570
Qy	501	GAAATATATGACACTATACGTTCGTTCGACGCTTATTTTGGGCTGTCTGTGGCCTGGTCT	560
Db	571	AGGCTCCATGACTATCCACACTTTGTGGCTTAATTCGGGCTCGTCTTTCCCGGGTCT	630
Qy	561	GCCAAAGCCTTACCCGAGGGAGCGGAGATTAAGATCAGACAGCAAGATCCAGTTT	620
Db	631	GTACAGGCCCACTGCGHAAAGACAGACACCGCAGGGCTTCGTCTACACTTACAGACT	690
Qy	621	GTCGCCATGCTGGGCGCCCTCTTCTTGTGATGTTCTGGCCAAATTTCAACTGTCTT	680
Db	691	CTTGCCATGATTTGGGACCATCTTCTGTGATCTTCTGGCTTACTTCAATGCTGCACT	750
Qy	681	GCTGAAAGTCCATTCGAAGAGAGATGCCGTGTTCAACACTTACTATGCTGAGCACT	740
Db	751	CACAGCGCTGGGGGCTGGGACAGCATCGGACGGCCCTCAACATCATCTCTCCCTGGCTGC	810
Qy	741	CAGGTGGTGAACGCACTCAAGGTCAATCTTGGCTCAACCCCAAGGGAATATACGAA	800
Db	811	CAGACCCCTTGGACCTTTGCTTGTACAGCCCTTTPAAGGGGAATGGGAGGCTTGACAT	870
Qy	801	GACTTATGTGACAGTGGGTGTTTGGACAGAGCGCTGCTGTGATCCTCGTCACT	860
Db	871	GCTCCACATCCAAATGACAGCGCTGGCTGGAGGGGTGTGTGTGGGGAACCTCAATGAAAT	930
Qy	861	GATCCCTTCTCCGTGGCTTGCATGTGTGGGTCTTGTGCTGGGAGCTGATCTCCGTGG	920
Db	931	GATGCTGACACCTTTTGGGGCTCTGGACAGCTGCTTCTTGGCTGGGAATGTCTCACGCT	990
Qy	921	GGGAGCCAGTACTGTCGGGGGTGTTTGAACCGAGTGTCTGGGGAATTTCCCAACAGCTCAT	980

APPLICANT:	Leite, Mario
APPLICANT:	Spytek, Kimberly A
APPLICANT:	Guo, Xiaojia (Sasha)
APPLICANT:	Fernandez, Elma
APPLICANT:	Li, Li
APPLICANT:	Kekuda, Rameesh
APPLICANT:	Liu, Xiahong
APPLICANT:	Casman, Stacie
APPLICANT:	Boldog, Ferenc
APPLICANT:	Paturajan, Meera
APPLICANT:	Blalock, Angela
APPLICANT:	Ballinger, Robert
APPLICANT:	Vernet, Corine
APPLICANT:	Tchernev, Vellizar T
APPLICANT:	Malysankar, Uriel M
APPLICANT:	Gusev, Vladimir
APPLICANT:	Rastelli, Luca
APPLICANT:	Mezes, Peter S
APPLICANT:	Ellerman, Karen
APPLICANT:	Heyes, Melvin P
APPLICANT:	Herman, John
APPLICANT:	Pena, Carol E A
APPLICANT:	Shankers, Richard A
APPLICANT:	Taupier Jr, Raymond J
APPLICANT:	Moore, No. US20040010119a11le
APPLICANT:	Shenoy, Suresh
APPLICANT:	Edinger, Shlomit
APPLICANT:	Gutcher, Erik
APPLICANT:	Stone, Dave
APPLICANT:	Millet, Isabelle
APPLICANT:	Peyman, John
APPLICANT:	Smithson, Glenda
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME	
FILE REFERENCE: 21402-269	
CURRENT APPLICATION NUMBER: US/10/074,978A	
CURRENT FILING DATE: 2003-01-07	
PRIOR APPLICATION NUMBER: 60/268,221	
PRIOR FILING DATE: 2001-02-12	
PRIOR APPLICATION NUMBER: 60/335,109	
PRIOR FILING DATE: 2001-10-31	
PRIOR APPLICATION NUMBER: 60/332,284	
PRIOR FILING DATE: 2001-08-14	
PRIOR APPLICATION NUMBER: 60/268,496	
PRIOR FILING DATE: 2001-02-13	
PRIOR APPLICATION NUMBER: 60/276,703	
PRIOR FILING DATE: 2001-03-16	
PRIOR APPLICATION NUMBER: 60/330,293	
PRIOR FILING DATE: 2001-10-18	
PRIOR APPLICATION NUMBER: 60/332,127	
PRIOR FILING DATE: 2001-11-21	
PRIOR APPLICATION NUMBER: 60/280,899	
PRIOR FILING DATE: 2001-04-02	
PRIOR APPLICATION NUMBER: 60/310,797	
PRIOR FILING DATE: 2001-08-08	
PRIOR APPLICATION NUMBER: 60/268,646	
PRIOR FILING DATE: 2001-02-14	
Remaining Prior Application data removed - See File Wrapper or PALM	

```

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver.
; SEQ ID NO 41
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-978A-41

```

Query Match	13.2%	Score 165.4	DB 17	Length 1351
Best Local Similarity	49.1%	Pred. No. 9.2e-42		
Matches 439	Conservative 0	Mismatches 456	Indels 0	Gaps 0

[illegible]

QY 26 TGTGGAGTGGGCAATTCCTGCTGGACGGCTTCCTAGCCAGTTCCTCTTGGAGAGTGT 320

Db 1055 CATCAGTGGGGCGCTCATGACGAGGCGTGGTCACTTCTTCAAGACCCCTAATCGT 996

QY 321 CATCACACTGTTCAATATTCGGCTGGGCACCATAGATGCTTGTGGTGCTATCTCAAT 380

Db 995 CGTGGCGGTGGAGAACTCATCAAGCTGACTTCTGCGTGGCCCTGTGCTGGCGGCTTT 936

QY 381 GAGTACTCTCTTTGGGGAAAGGCACCTTGGGGCAGTTGGTGGTGCATGCTGCTGGTGAAGT 440
 Db 935 TGGGGCAGTTCTTGGGTAAAGTCACGCCCATTCACGCTGCTCATCATGACTCTTCTTCCAAAGT 876
 QY 441 GACACGTTTAAAGCACTGAGAGTGTGATCATGATTAATCTTCAACACAGACTACACAT 500
 Db 875 GACCTCTTGGCTGTGATGAGTTCAATCTCCCTTAAGCTGCTAAAGGTAAAGGATGAG 876

[illegible]

621 GTCTGCATGCTGGGCGCCCTCTTCTTGTGAGATGTTCTGGCCAGTTTCAACTGCTCT 680
622 |||||
623 |||||
624 |||||
625 |||||
626 |||||
627 |||||
628 |||||
629 |||||
630 |||||
631 |||||
632 |||||
633 |||||
634 |||||
635 ATCTTACCAATGGAGATGCGCTCTTCCCTCTTCTGAGATGATCGCCAGCTTCAACTGAGCAT 636
636 |||||
637 |||||
638 |||||
639 |||||
640 |||||
641 |||||
642 |||||
643 |||||
644 |||||
645 |||||
646 |||||
647 |||||
648 |||||
649 |||||
650 |||||
651 GCTGAGAGTCCATGGAAGAAGATGCCCTGTTCAACACTTACTGTGTCGACGT 740
652 |||||
653 |||||
654 |||||
655 |||||
656 |||||
657 |||||
658 |||||
659 |||||
660 |||||
661 |||||
662 |||||
663 |||||
664 |||||
665 |||||
666 |||||
667 |||||
668 |||||
669 |||||
670 |||||
671 |||||
672 |||||
673 |||||
674 |||||
675 |||||
676 |||||
677 |||||
678 |||||
679 |||||
680 |||||
681 |||||
682 |||||
683 |||||
684 |||||
685 |||||
686 |||||
687 |||||
688 |||||
689 |||||
690 |||||
691 |||||
692 |||||
693 |||||
694 |||||
695 |||||
696 |||||
697 |||||
698 |||||
699 |||||
700 |||||

741 CAGCGTGTGACAGCCATCTCAGGGGTCCTTGTGCTCACCCCAAGGAAGATCAGCA 800
575 CTGGTGCTTACCTCGGTGGCAATATCCAGTCCCTGCACAAGAAAGGCAAGCTGCAGAT 516
801 GACTTATGTGCACAGTGCAGGTGTGGCAGAGAGCGTGGCTGTGGTACCTCGTCACT 860

Y	861	GATCCCTTCCTCCGTGCTTCAGATGCTGCTGGCTCTTGTGGCTGGGCTGATCTCCGCTGG	920
b	455	GATGCTCATGCTTACGGTGCCCTATATGAGGCTTCGTCTGCGGCATCTATCTCCACCT	396
Y	921	GGGAGCAGATACCTCCGGGGTGTGTGAACGAGTGTGGGAGTTCCACAGCTTCAT	980

b

```

981 CATGGGCTTCAAACTTCACCTTGCTGGGTCTCTCGTAGAGATCATCTAATTGTG   1035
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
335 CATTAAACATCTGCATGCGCATTCTCGGCATCATGGGGGATCGTGGAGCTGTGTG   281

```

```

RESULT 5
US-09-949-145-5
; Sequence 5, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glycop
; FILE REFERENCE: Docket 454-11
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-949-145-5

```

Query Match	13.2%;	Score 165.4;	DB 9;	Length 1440;
Best Local Similarity	49.1%;	Pred. No. 9.5e-42;		
Matches 439; Conservative	0;	Mismatches 456.	Indels 0.	Canc 0.

[illegible][illegible]

QY 381 GGATGCTGTCTTGGGGGAAGTCAACTTGGGCGAGTTGGTGAGTGTCTGTGGAGGT 440
Db 417 TGGGGCAGTTCTGGGTAAAGTCAGCCCCAATCAGCTGCTCATCAAGACTTCTTCCAAGT 476
QY 441 GACAGCTTTTGGCACTGAGATGGTGCATCGTAAATCTTCAACAGACTACCAAT 500

[illegible][illegible]

117 ATCTTACCAATGGGAGACAGCCAGACCGACCCATCAACCTTACTGTCTCTTTGGAGG 776

741 CACGCGTGTGACAGCATCTCTCAGGGTGCATCCTTGGCTACCCCCAAGGAAGATTCAGCAA 800

777 CTGCGTGCTTACTCGGTGGCAATTTCCATGAGCCCTTGCACAAAGAGGGCAAGGCTGCAGCAT 836

801 GACTTATGTGCACAGTGGCGGTGTGGCAGAGAGCGTGGCTGTGGGATACCTCGTGTACCTT 860

837 GGAGCATCATCGAAGATGCCACGCTCCGACAGAGGGTGGCCCTGGATACCGCTGTAGAT 896
861 GATCCCTTCTCCGTGGCTTGGCAATGGTGGCTTGTGGCTGGGCTGATCTCCGTGG 920

Db 897 GATGCTCATGCTTACGCGTCCCTCATCATCGGCTTGTGCGCGCATCATCTCCACCT 956
Qy 921 GGGAGCCAAAGTACCTGCGGGGGTGTGTAACCGAGTGTGGGATTCGCCACAGCTCAT 980
Db 957 GGGTTTGTATACCTGACCCCATTCCTGGAAGTCCCGGCTGCATCTCAGACACATGTGG 1016
Qy 981 CATGGGCTAACATTCCTGAGTGTGCTGCTGCTGAGAGATCATCTACATTTGTG 1035
Db 1017 CATTAACATCTGATGAGCATTCCTGATCATATGAGGGGATCGTGGGTGTGTG 1071

RESULT 6

US-09-949-145-1
; Sequence 1, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193809
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(1952)
US-09-949-145-1

Query Match 13.2%; Score 165.4; DB 9; Length 1952;
Best Local Similarity 49.1%; Pred. No. 1.1e-41;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAATCTGACCGTGTATGCGGCGCATTTGGGCTTCTTCACCTC 200
Db 201 CTACCCAGCTTCCAGACGCGTGAATGCTTGTGGGCTTCCGCTTCATGAC 260
Qy 201 GAGTTCCGGAGACACAGCTGAGAGTGGGCTTCAACCTCTTACGTGGGCTTGG 260
Db 261 TTTCCTCAGCGCTACGCGCTTCAAGCGCGGCGCTTCACTTCTGTGGAGCGCTTGG 320
Qy 261 TGTGCAATGGGCAATCCTGCTGACGCGCTTCTGAGCGAGTTCCTTCTGGGAAGTGT 320
Db 321 CATCCATGGGCGGTGCTCATGAGGCGTGTCCACTTCTTAAAGACCGTACATCGT 380
Qy 321 CATCACTGTTCAATTCGCGTGGCCACCATGAGTGTCTTGTGGGTGTGATCTGAGT 380
Db 381 CGTGGCGGTGAGAACTCATCAACGCTGACTTCTGGGCGCTCTGTCTGCGGCTT 440
Qy 381 GGAATGCTGTGGGGAAGTCACTTGGCGCACTTGTGTGATGTGCTGTGAGAGT 440
Db 441 TGGGGCGTTCTGGTAAAGTCAAGCCCATTCACCTCTCATGAGCTTCTTCCAGT 500
Qy 441 GACAGCTTTAGGCAACCTGAGAGTGTATCATGATTAATCTTCAACACAGACTACCAT 500
Db 501 GACCCCTTGTGCTGTGATGATGATTCATTCCTTCACTGTGTAAGTGAAGATGACAG 560
Qy 501 GAACATATGACATCTACGTGTGTGACGCTATTTTGGGCTGTGTGTGCTGTGCTGT 560
Db 561 AGGCTTCATGACATTCACATTTGAGGCGCTTACCTTGGGCTCACAGTGAACCGGATCT 620
Qy 561 GCCAAAGCTTACCGAGGGAAGGAGTAAGATGACACAGACAGATACCATGTTT 620
Db 621 CTACCGAGGCAACTGAGACAGAGAGAGACAGATTTCTGTACAGTGGAGCT 680
Qy 621 GTCTGCATGCTGGGCGCTCTTCTGTGTGATGTCTGCGCAAGTTTCAACTGTGCTCT 680

Db 681 CTTGCGATATTTGGACCCCTTCTGTGATGTACTGTGCGCCAGCTTCAACTCAGCCAT 740
Qy 681 GCTGAGAAATCCATGCAAGAAAGAAATGCGGTGTTCACACCTACTATGCTGAGAGT 740
Db 741 ATCTTACATAGGGAACGACGACGACGAGCGCCATCAACACTTACTGTCTTGGGAGC 800
Qy 741 CAGGTGTGACAGCCATCTCAAGGTGATCTTGGCTCACCCCGAGAGATCAGCA 800
Db 801 CTGGGTCTTACCTCGGTGGCAATATCAGATGCGCTGCACAGAAAGGCAAGTGCAT 860
Qy 801 GACTTATGTGACAGTTCGGGTGTGGCAGAGCGGTGCTGTGGGTACCTGTCTACCT 860
Db 861 GGTGCATCATCAGATTCACAGCTGCGAGAGGGGTGCGGTGGGTACCGTGTGAGAT 920
Qy 861 GATCCCTTCCGCGGTTCGATGCGTGGGTGCTGTGGGTGGGTGATCTCGTGG 920
Db 921 GATGCTCATGCTTACGCTGCTCATCATGAGTCTTGTGTGCGGCAATCTCACCCT 980
Qy 921 GGGAGCCAAAGTACCTGCGGGGATGTGTAACCGAGTCTGGGGAATCCGACAGCTCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGATCCCGGCTGCATCCAGACACATGTGG 1040
Qy 981 CATGGGCTAACATTCCTGAGTGTGCTGTGCTGTGAGAGATCATCTACATTTGTG 1035
Db 1041 CATTAACATCTGATGAGCATTCCTGATCATATGAGCGGATCGTGGGTGTGTG 1095

RESULT 7

US-10-172-118-1635
; Sequence 1635, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1635
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016321
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1635

Query Match 13.2%; Score 165.4; DB 17; Length 1952;
Best Local Similarity 49.1%; Pred. No. 1.1e-41;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAATCTGACCGTGTATGCGGCGCATTTGGGCTTCTTCACCTC 200
Db 201 CTACCCAGCTTCCAGACGCTGACGATGATGTCTTGTGGGCTTCCGCTTCCTCATGAC 260
Qy 201 GAGTTCCGGAGACACAGCTGAGAGTGGGCTTCAACCTCTTCACTGTGGGCTTGG 260
Db 261 TTTCCTCAGCGCTACGCGTTTCAAGCGCGGCTTCACTTCTGTGGAGCGCTTGG 320
Qy 261 TGTGCAATGGGCAATCTGCTGACGCGCTTCTGAGCAAGTTCCTTCTGTGGAAGTGT 320
Db 321 CATCAAGTGGGCGGTGCTCATGAGGCGGTGCTTCACTTCTTAAAGACCGCTACATCGT 380

Sequence 6, Application US/09949145
Patent No. US20020055622A1
GENERAL INFORMATION:
APPLICANT: New York Blood Center
TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
FILE REFERENCE: Docket 454-31
CURRENT APPLICATION NUMBER: US/09/949, 145
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/230660
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 1497
TYPE: DNA
ORGANISM: Mus musculus
US-09-949-145-6

Query Match 12.0%; Score 150.2; DB 9; Length 1497;
Best Local Similarity 47.5%; Pred. No. 7.4e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

97 CACTATATACGCTTCTTATAGAGATCAAAAGGCGCTCGTGCGATCTATCAAGTTGGCCAA 156
136 CGCAAGAACATCTCCAGCGAGCTTGAGACGAGTTTACTATCGCTACCGAGCTTCAG 195
157 GATCTGACCGGTGATGGCGCATGGCTGGCTTCCCTCACTCGAGTTTCCGAGACAC 216
196 GATGTACACCGCATGCTCTTGCTGGGCTTCCCTCATGACCTTCTCGACGCTAC 255
217 AGCTGAGCAGTGTGGCTTCAACCTTTCATGCTGCGCTGGTGTGTCAGTGGCAATC 276
256 GGCCTTCAGCGCTGAGGCTTCACTTCTGCTGCGACCTTTCGCGATCAAGTGGGCACTG 315
277 CTGCTGACGCGCTTCTCGACGCACTTCCCTTGGGAAAGTGTGATCACTGTTCACT 336
316 CTATGACGAGGATGCTTCTTACTTGTGAAGAAAGCCACATGTTCTGAGCGCTGAGAAC 375
337 ATTCGCGTGGCGACATGAGTGTGCTTGTGCGGTGCTGATCTCAGTGAAGTGTGCTGGGG 396
376 ATCATTCAGAGTACCTTCTGTGTGCACTTCTGCTGTGCGCTTCCGGGGCATTTCTAGGC 435
397 AAGTCAACTTGGCGCAGTTGTGTGTGATGCTGTGCTGTGAGTGAAGCTTTTAAAGCAAC 456
436 AAGGTACCGCGATGACGCTGCTCATTTATGACCTTCTTCCAAAGGACCTCTTCAAGTG 495
457 CTGAGATGCTCATCAATATCTTCAACACAGACTACCATGAACATGATGACATC 516
496 AATGAGTTCACTCTCTGAACTGTATGAGGCAAGATGACGGGGCTCTATGACATC 555
517 TACGTGTGCGACCTATTTTGGGCTGTGTGCTGCGCTGGTGGCCCAAGCCCTTACCC 576
556 CACACATTTGGCGCTTACTTGTGGCTCAAGTGAAGCTGATCTCTAACGAAAAACCTG 615
577 GAGGGAACGAGAGATTAAGATCAGACAGCAAGATACCAATTGTGTGCCATCTGGGC 636
616 GATCAGAGCAAGAGAGACAGAGCTCACTGTAACACATCGACCTTTGCGCATATTGGC 675
637 GCCCTCTTTCTTGTGATGTTTGGCCAAATTTCATCTGTCTCTGCTGAGAGTCAATC 696
676 ACCCTCTTCTTGTGATTAATCTGGCCAGTTTCAATTGACAGTCTCTTCCAGGAGAT 735
697 GAAAGGAAGATGCGGTTCACACCTAATGCTGTAGACAGTCAAGGTGTGTAGACGCC 756
736 GCCCAGACCGAGAGCCCTTAATACCTAATCTTCTTGGGAGGAGGTGCTTAACACA 795
757 ATCTCAGGTCATCTTGGCTCAACCCCAAGGAAGATCAAGCAAGCTTATGTGACAGT 816
796 GTGACATATATCAATATTGTACACAAGAGGCAAGTTGATATGTGTGCAATCAAGAT 855
817 GCGGTGTGGCAGAGCGGT 876
856 GCCACGCTTGCAGGT 915

877 CTTCGATGT 936
916 GGGGCTTCATCGTGGGGTCTTCTGCGGCAATTTTCCACCCTAGGATTTGCACTACTA 975
937 CCGGGGTGTGTATACCAAGT 996
976 ACAGCATCTCTGAGT 1035
997 AGCTGCTGGGTGCTGT 1074
1036 GGCATTCCTGCGCATATAGGCGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1074

RESULT 10

US-09-949-145-2
Sequence 2, Application US/09949145
Patent No. US20020055622A1
GENERAL INFORMATION:
APPLICANT: New York Blood Center
TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
FILE REFERENCE: Docket 454-31
CURRENT APPLICATION NUMBER: US/09/949, 145
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/230660
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2097
TYPE: DNA
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AF193810
DATABASE ENTRY DATE: 1999-12-22
RELEVANT RESIDUES: (1)..(2097)
US-09-949-145-2

Query Match 12.0%; Score 150.2; DB 9; Length 2097;
Best Local Similarity 47.5%; Pred. No. 8.8e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

97 CACTATACGCTTCTTATAGAGATCAAAAGGCGCTCGTGCGATCTATCAAGTTGGCCAA 156
258 CGCAAGAACATCTCCAGCGAGCTTGAGACGAGTTTACTATCGCTACCGAGCTTCAG 317
157 GATGTACCGGTGATGGCGCATTTGGCTTCCCTCACTGAGATTTCCGAGACAC 216
318 GATGTACACCGCATGCTCTTGTGTGCGCTTCCCTCATGACCTTCTGACGCGTAC 377
217 AGCTGAGCAGTGTGGCTTCAACCTTTCATGCTGCGCTGTGTGTCAGTGGGCAATC 276
378 GGCCTTCAGCGCTGAGGCTTCACTTCTGCTGCAAGCTTTCGCGATCAAGTGGGCACTG 437
277 CTGCTGACGCGCTTCTGAGCAGTTCCTTGGGAAAGTGTGATCACTGCTTCACT 336
438 CTATGACGAGGATGTTCACTTACTTGAAGAAAGCCACATTTCTGAGGCTGAGAAC 497
337 ATTGCGTGGCCACATGAGT 396
498 ATCATTCAGCTGACTTCTGTGTGCACTTCTGCTGTGTGCGCTTCCGAGGAGTTCTAAGC 557
397 AAGTCAACTTGGGCGAGTTGT 456
558 AAGGTACCGCGATGACGCTGCTCATTTATGACCTTCTTCCAAAGTACTCTTTACAGTG 617
457 CTGAGATGTCATCAATATCTTCAACACAGACTACCATGAACATGATGATGACATC 516
618 AATGATTCATCTCTGAACTGTATGAGCAAGATGACGAGGCGCTTATGACATC 677
517 TACGTGTGCGACCTTATTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
678 CACACATTTGGCGCTTACTTGTGGCTCAAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 737

QY 577 GAGGGAACGAGAGATTAAGATCAGACAGCAACATACCAGTTGTGTCATGCTGGGC 636
DB 738 GATCAGAGCAACAGACAGACAGCTCAGTGTACCACTGGACCTTTTGGCATTGTGGC 797
QY 637 GGCCTCTCTTGTGATGTGTCGCAAGTTTCAACTCTGCTCTGCTAGAGATGCAATC 696
DB 798 ACCCTCTCTTGTGATGTGTCGCAAGTTTCAACTCTGCTCTGCTAGAGATGCAATC 857
QY 697 GAAAGGAAGATCCGCTGTTCAACACTATGCTGTAGACAGTGGGTGTGACAGCC 756
DB 858 GCCCAGCAACGAGCAGCCCTCAATCTCTCTGCGAGCGAGTGTCTAACACA 917
QY 757 ATCTAGAGTCACTCTTGGCTTCAACCCCAAGGAAAGATCAGCAAGATTTATGACAGT 816
DB 918 GTGACAGTATCAGATATGTACACAAAGAGGCAAGTTGATGTGTGCATCATCAGAT 977
QY 817 GCGGTGTGGCAGAGGCGTGGCTGTGGTACCTGCTGACCTGATCCCTTCTCGTGG 876
DB 978 GCCACGCTTGCAGAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1037
QY 877 CTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
DB 1038 GACCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
QY 937 CCGGGGTGTGTAAACAGTGTCTGGGATTTCCCAAGCTTCATCATGAGGCTACACTTC 996
DB 1098 ACGCATTCCTGAGTCCCGCTTCGATCCAGACACATGTGCTATTCACCAACTGCTGAC 1157
QY 997 AGCTTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
DB 1158 GGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1196

RESULT 11
US-10-047-730/c
; Sequence 730, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 730
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-730

Query Match 11.9%; Score 148.8; DB 17; Length 2415;
Best Local Similarity 65.9%; Pred. No. 2,7e-36;
Matches 216; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 758 TCTCAGAGTATCTTGGCTCAACCCCAAGGAAATCAGCAAGACTTATGTGCAAGTG 817
DB 1875 TCTCTTACCTTGTCTCTTACCCACAGCTATTTCTTTCAGAGACTTATGTGCAAGTG 1816
QY 818 CCGTGTGGCAGAGGCGTGGCTGTGGGTACCTCGGTGTACCTGATCCCTTCTCGGTGGC 877
DB 1815 CCGTGTGGCAGAGGCGTGGCTGTGGGTACCTCGGTGTACCTGATCCCTTCTCGGTGGC 1756
QY 878 TTGCAATGTGCTGGTCTTGTGCTGGCTGTATCTCGTGGGGAAGCAAGTACTGCG 937
DB 1755 TTGCAATGTGCTGGTCTTGTGCTGGCTGTATCTCGTGGGGAAGCAAGTACTGCG 1696
QY 938 CCGGGGTGTGTAAACAGTGTCTGGGATTTCCCAAGCTTCATCATGAGGCTACACTTCA 997
DB 1695 CCGTAAAGAACTAGCAACTAATCTCTCTGCTTGTGGTGAAGGCGACAGAGACGCTGGG 1636

QY 998 GCTTGTGGGTCTGCTTGGAGAGATCATCTTACTTGTGCTGCTGCTGCTGCTGCTGCTG 1057
DB 1635 ACCGTATGGGCGCACTGTGCAATGACAGCTGATGTAGGAGGTGTGGGCAATTCCTTA 1576
QY 1058 GAGCCGCAATGAGCATGATTTGGCTTCA 1085
DB 1575 TTGCTTCAACGCTTGTGTAGGAGATCCA 1548

RESULT 12
US-10-027-632-282039
; Sequence 282039, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282039
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282039

Query Match 11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 2,7e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 758 TCTCAGAGTATCTTGGCTCAACCCCAAGGAAATCAGCAAGACTTATGTGCAAGTG 817
DB 241 TCTCTTACCTTGTCTCTTACCCACAGCTATTTCTTTCAGAGACTTATGTGCAAGTG 300
QY 818 CCGTGTGGCAGAGGCGTGGCTGTGGGTACCTCGGTGTACCTGATCCCTTCTCGGTGGC 877
DB 301 CCGTGTGGCAGAGGCGTGGCTGTGGGTACCTCGGTGTACCTGATCCCTTCTCGGTGGC 360
QY 878 TTGCAATGTGCTGGTCTTGTGCTGGCTGTATCTCGTGGGGAAGCAAGTACTGCG 937
DB 361 TTGCAATGTGCTGGTCTTGTGCTGGCTGTATCTCGTGGGGAAGCAAGTACTGCG 420
QY 938 CCG 940
DB 421 CCG 423

RESULT 13
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040
```

```
Query Match 11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 2,7e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 758 TCTCAGGATATCTTGGCTACCCCAAGGAGATCAGACACTTATGTGACAGTG 817
    |||
DB 241 TCTCTACCTTCTTCTTACCAACGCTATTTCTTGACACTTATGTGACAGTG 300
    |||
QY 818 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 877
    |||
DB 301 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 360
    |||
QY 878 TTGCGATGCTGCTGGCTTTGTGGCTGGCTGATCTCCGTCGGGGAGCCAACTACTGC 937
    |||
DB 361 TTGCGATGCTGCTGGCTTTGTGGCTGGCTGATCTCCGTCGGGGAGCCAACTACTGC 420
    |||
QY 938 CGG 940
    |||
DB 421 CGG 423
```

```
RESULT 14
US-10-027-632-282039
; Sequence 282039, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282039
```

```
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282039
```

```
Query Match 11.5%; Score 144.6; DB 17; Length 505;
Best Local Similarity 86.9%; Pred. No. 2,7e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 758 TCTCAGGATATCTTGGCTACCCCAAGGAGATCAGACACTTATGTGACAGTG 817
    |||
DB 241 TCTCTACCTTCTTCTTACCAACGCTATTTCTTGACACTTATGTGACAGTG 300
    |||
QY 818 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 877
    |||
DB 301 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 360
    |||
QY 878 TTGCGATGCTGCTGGCTTTGTGGCTGGCTGATCTCCGTCGGGGAGCCAACTACTGC 937
    |||
DB 361 TTGCGATGCTGCTGGCTTTGTGGCTGGCTGATCTCCGTCGGGGAGCCAACTACTGC 420
    |||
QY 938 CGG 940
    |||
DB 421 CGG 423
```

```
RESULT 15
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282040
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282040
```

```
Query Match 11.5%; Score 144.6; DB 17; Length 505;
Best Local Similarity 86.9%; Pred. No. 2,7e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 758 TCTCAGGATATCTTGGCTACCCCAAGGAGATCAGACACTTATGTGACAGTG 817
    |||
DB 241 TCTCTACCTTCTTCTTACCAACGCTATTTCTTGACACTTATGTGACAGTG 300
    |||
QY 818 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 877
    |||
DB 301 CGGTGTTGGCAGAGGCGGTGGTGGTACCTGCTGACCTGATCCCTTCCGGGC 360
    |||
QY 878 TTGCGATGCTGCTGGCTTTGTGGCTGGCTGATCTCCGTCGGGGAGCCAACTACTGC 937
    |||
```

Db	361	TTGCCATGGTGGCTGGGCTTTGTGGCTGGGCTGATCTCCGTGGGGGAGCCAACTACTGC	420
Qy	938	CGG	940
Db	421	CGG	423

Search completed: March 23, 2005, 07:24:10
Job time : 4489 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 23:43:09 ; Search time 3942 seconds
(without alignments)
12108.725 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctcctaacgaccgcg.....attgctgttgattctaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: gb_ests1.*
2: gb_ests2.*
3: gb_hlc.*
4: gb_ests3.*
5: gb_ests4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	824.4	65.7	964	5	BX406635
2	633.8	50.5	685	5	BU664461
3	629.8	50.2	676	5	BX106517
4	617.2	49.2	685	5	BU655978
5	602.2	48.0	1457	3	AK079335
6	597.4	47.6	706	5	BU661907
7	585.4	46.7	878	5	BX406634
8	560.4	44.7	631	5	BU662207
9	559	44.6	632	5	BU665450
10	552	44.0	613	5	BU664659
11	548.8	43.8	888	7	CK453280
12	542.4	43.3	607	5	BU657533
13	498.6	39.8	975	3	AK089642
14	429.2	34.2	458	6	CD627633
15	427.2	34.1	612	5	BU663660
16	418.6	33.4	862	7	CK775698
17	410	32.7	461	6	CD627634
18	405.6	32.3	587	5	BU661598
19	399.2	31.8	480	5	BU659596
20	398.6	31.8	683	4	BM719724
21	392.4	31.3	443	1	AL699045
22	377.8	30.1	873	7	CN166957
23	375.6	30.0	870	7	CK452223
24	366	29.2	804	7	CO807980

25	359	28.6	429	5	BU657883	BU657883 c129d10.z
26	349.2	27.8	721	6	BY750607	BY750607 BY750607
27	343.2	27.4	551	4	BM254191	BM254191 515552.MA
28	321.2	25.6	660	6	BY722497	BY722497 BY722497
29	317.6	25.3	643	4	BM684087	BM684087 UI-E-RT1-
30	314.4	25.1	701	6	BY734317	BY734317 BY734317
31	302.2	24.1	434	7	NS9044	NS9044 yv60b12.r1
32	299.6	23.9	375	5	BU664580	BU664580 c1119p02.r1
33	294.2	23.5	499	2	BF603905	BF603905 269531.MA
34	288.2	23.0	488	2	BF602079	BF602079 267224.MA
35	280.4	22.4	550	9	CG573938	CG573938 OSTR206200
36	273.6	21.8	515	5	BX529358	BX529358 BX529358
37	271.2	21.6	458	4	BG944752	BG944752 ax54d09.x
38	270.4	21.6	416	7	R10548	R10548 yf31e01.r1
39	257.8	20.6	601	5	BU661774	BU661774 c116g01.z
40	255.8	20.4	522	9	CG634208	CG634208 OSTR354881
41	253.6	20.2	343	7	T84327	T84327 yd37a05.r1
42	250	19.9	304	9	CG669296	CG669296 OSTR466105
43	243	19.4	515	2	BF191605	BF191605 239237.MA
44	240.8	19.2	474	6	BY563146	BY563146 BY563146
45	240.8	19.2	476	6	BY559962	BY559962 BY559962

ALIGNMENTS

RESULT 1
BX406635
LOCUS
DEFINITION BX406635 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM013YA17 5-PRIME, mRNA sequence.
BX406635
BX406635.2 GI:46952745
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 964)
I.J.M.B., Gruber, C., Jessee, J., and Polayres, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30762708.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 854.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0AM013AA09QP1c=854.r.
FEATURES
Location/Qualifiers
1..964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM013YA17"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: Liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 65.7%; Score 824.4; DB 5; Length 964;
Best Local Similarity 93.0%; Pred. No. 6.5e-216;

TITLE FANTOM Consortium.
JOURNAL Functional annotation of a full-length mouse cDNA collection
REFERENCE Nature 409, 685-690 (2001)
AUTHORS
TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL Group Phase I & II Team.
REFERENCE Analysis of the mouse transcriptome based on functional annotation
AUTHORS of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1457)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carrinci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Onno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.jp,
URL: http://genome.gs.c.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gs.c.riken.jp/
URL: http://fantom.gs.c.riken.jp/.
FEATURES
source
Location/Qualifiers
1. 1457
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9630028A03"
/db_xref="taxon:10090"
/clone="9630028A03"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
10. 1266
/note="unnamed protein product: Rhesus blood group CE and
D (MGD|MGJ:1202882, GBI|NM_011270, evidence: BLASTX, 99%,
match=1257)
putative"
/codon_start=1
/protein_id="BAC37610.1"
/db_xref="GI:26347923"
/translation="MSGKYPRSRRCCLPLMALEIQTRFSLFCEPIPHDTAYDHRPM
ASTQVKNLTLMALGFGFLSSFSRRHSWSVAFNFLMLAGVQGITLLDHFQGVQ
WKININLSIQIATMTLPVLISAGAVLGYNLVOLTVNMLMAMAGARFADKVF
KMTKEHIMGHVFGAFGLTAVMWLSRSLPRRGENAQTCKQMAKSSLSLPMGLTL
FLVIFWPAIVASLLEGTKKRNAVNTYALASVPTASATSSHPGKLNMYHNA
VLAGVAIVAGGLISGSPWISGLIGLISMGACSPACSNHINONSSGIHYTNG
LPELGLATYYCQIVTERKSSDLMITTVTVTHGALSPFVAMGMVTGLTGLLSVR
VMPAPRAKTFDDQTFWEPFLAVGF"
1436. 1441
/note="putative"
1457
/note="putative"
ORIGIN
Query Match 48.0%; Score 602.2; DB 3; Length 1457;
Best Local Similarity 69.7%; Pred. No. 1.6e-154;

Matches 884; Conservative 0; Mismatches 358; Indels 27; Gaps 4;
QY 1 ATGAGCTCTAAGTACCGCGCTGTCGCGGCTGCTGCGCCCTGAGCCCTTAACACTG 60
DB 10 ATGGGCTTAAGTACCGCGCTGTCGCGGCTGCTGCGCCCTGAGCCCTTAAGACTA 69
QY 61 GAAGAGCTCTAATCT 120
DB 70 CAGACAGCTTTTATGCT 129
QY 121 CAAAAGGGCTGCTGCGATCTCTAAGTTGGCCAGATCTGACCCGATGAGCGGCATT 180
DB 130 CACA---GGTTCAATGGGAGACTATCAAGTCTCCGAATTTGACCTTCATGGACGCTTGG 186
QY 181 GGGTGGGCTCTCTCACTCTGAGTTTCCGAGACACAGCTGAGAGTGGCCTTCAAC 240
DB 187 GGGCTTGGCT 246
QY 241 CTCTTCATGCTGCGCTTGGTGTGCACTGGGCAATCTGCTGAGCGGCTTCTGAGCCAG 300
DB 247 CTCTTCATGTTGGCCCTCGGGGTGCAAGGAAACAATCTGCTGAGCAATTTCTGAGCCAG 306
QY 301 TTGCTCTTGGAGAGTGTGATCACTGTTCAATTTGGGCTGGCCACCATGATGCT 360
DB 307 GTCTCCATGGAACAAGATCAACATCTGTCCAGATTCAGATGACTACATGAGCACC 366
QY 361 TTGTGCGTGTGATCTCAGTGAATGCTGTGGGGAAGGTCAACTTGGGCACTTGCTG 420
DB 367 TTACTGTGTGATCTCAGCGGGGCTGTCTGGGGAAGGTCAACTGTGTGCACTTACC 426
QY 421 GTGATGCTGTGAGTGAAGTGAAGCTTTAGGCACTGAGATGTCTATCAGTAATATC 480
DB 427 GTATGATGCTGATGAGGGAAGTGGCTTTGATGATCAGATTTGGCGAGAGAGTTC 486
QY 481 TTCAACACAGACTACACATGAACATGATGACATCTACTGTTCGACCTATTTGGG 540
DB 487 TTCAAAATGACAGAAACATCATCATGATGACAGGGAACGTGTTGGGCTATTTTGGG 546
QY 541 CTGTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 547 CTAACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
QY 601 ACA-----GCAGATTAACCAAGCTTGTCTGCAATGCTGAGGCGCCCTCTTC 645
DB 607 ACAGAGAGGTGAATAGGCTAAGAGCTCAAGCTGTGTTCCATGCTGGGACCTCTTT 666
QY 646 TTGTGATGTTCTGCGCAAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
DB 667 TTGTGATATTTCTGCGCAGCTATCACTGCTCTCTCTGGAAGGGAAGAAAG---G 723
QY 706 AATGCGGTGTTCAACCTCTATGCTGATGAGTCAAGTCAAGGTCAGTCAAGCATCTCAGG 765
DB 724 AATGCTGTGTTCAACCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 783
QY 766 TCATCTGTGCTCAACCCCAAGGGAAGATGAGAAAGATTATGTGACAGTGCCTGTTG 825
DB 784 TCAGCCTGAGTCAACCTCAAGGGAAGATGAATGATGTTCAATTCACATGACAGTCTG 843
QY 826 GCAAGAGCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
DB 844 GCAAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
QY 886 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
DB 904 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
QY 946 TGTAACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
DB 964 TTGAACCAATGCTGAGAACTCAAGTGG-----GATCACTACACACTTGGCTTGGCCG 1017
QY 1006 GGTCTGCTGGAAGATCATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
DB 1018 GGTCTGCTGGAACACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 1077

QY 1066 AATGGCATATGCTTCCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGG 1125
 DB 1078 GATCTCTGATCATCAACCCAGACGCTCATCATCTGGGCTCTACGCTTCCCTGTGGG 1137
 QY 1126 ATAGCTCTACGCTGTCTCTCTGACAGGTTGCTCTTAATCTTAAATATGAAAGCA 1185
 DB 1138 ATGGGATGATGATCTGACCTCTCTCAGGTTGCTCTTAAGTGTGAGAGGCT 1197
 QY 1186 CCTCATGAGGCTAAATATTTGATGACCAAGTTTCTGAACTTCTCATTTGGCTGT 1245
 DB 1198 CCCCGTCGCGCAAGATTTTATGATCATGACCTTCTGGAGTTCCACACTTGGCGGT 1257
 QY 1246 GGATTTTAA 1254
 DB 1258 GGATTTTAA 1266

RESULT 6
 BU661907
 LOCUS BU661907 706 bp mRNA linear EST 30-SEP-2002
 DEFINITION c178e10.21 Hembase; Erythroid Precursor Cells (LCB:c1 library) Homo
 sapiens cDNA clone c178e10 5', mRNA sequence.
 BU661907
 ACCESSION BU661907.1 GI:23374089
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
 Gene Expression in Human Erythroid Precursor Cells
 Unpublished (2002)
 Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jlm7f@nih.gov
 The 'c1' library was constructed by Alexander Gubin, Ph.D. in the
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
 analyses by National Institutes of Health Intramural Sequencing
 Center (NISC). More information available at:
 http://hembase.nidk.nih.gov
 Plate: 78 row: e column: 10
 Seq primer: 5' lambda-TripLex2 Sequencing Primer.
 Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="c178e10"
 /sex="unknown"
 /issue_type="blood"
 /cell_type="Erythroid Precursor Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_host="DH5alpha"
 /lab_host="Hembase; Erythroid Precursor Cells (LCB:c1
 library)"
 /note="Organ: blood; Vector: pTriplex2; Site: 1: Sf11;
 Site: 2: Sf11; A complementary DNA (cDNA) library from
 human erythroid precursor cells was constructed using
 SMART PCR (polymerase chain reaction) cDNA library
 Construction Kit (Clontech, Palo Alto, CA) according to
 the manufacturer's directions, but with slight
 modifications. Briefly, reverse transcription was
 performed in the presence of 1 umol/L peptide nucleic acid
 (PNA) oligos

(N-terminal)-biotin-GTC-CAC-CCG-AAA-GTT-G-(C-terminal) and
 (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
 . Synthesized cDNA was digested with Sf11 and
 size-selected on a 1% agarose gel (>800bp). Large-scale
 sequencing of the library was performed by the NIH
 Intramural Sequencing Center (NISC).
 Http://www.nisc.nih.gov/."

ORIGIN

Query Match 47.6%; Score 597.4; DB 5; Length 706;
 Best Local Similarity 97.4%; Pred. No. 2,6e-153;
 Matches 607; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATAGGCTCTAAGTACCCCGGCTGTGTCCGCGCTGCTGCGCCCTTGGGCTTAACATG 60
 DB 84 ATAGGCTCTAAGTACCCCGGCTGTGTCCGCGCTGCTGCGCCCTTGGGCTTAACATG 143
 QY 61 GAAGCAGCTCTCATTTCTCTCTTATTTTATACCACTAAGAGCTTCTTAAGAGAT 120
 DB 144 GAAGCAGCTCTCATTTCTCTCTTATTTTATACCACTAAGAGCTTCTTAAGAGAT 203
 QY 121 CAAAGGCGCTGTGCGATCTTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCAT 180
 DB 204 CAAAGGCGCTGTGCGATCTTATCAAGTTGGCCAAAGATCTGACCGTGATGGCGCAT 263
 QY 181 GGCTTGGGCTTCTCACTCCAGCTTCCGAGACACACAGTGGAGAGATGTGGCTTCAAC 240
 DB 264 GGCTTGGGCTTCTCACTCCAGCTTCCGAGACACACAGTGGAGAGATGTGGCTTCAAC 323
 QY 241 CTCTTCATGCTGCGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGACAG 300
 DB 324 CTCTTCATGCTGCGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGACAG 383
 QY 301 TTCCCTTCTGGGAAGTGTGTATCACTGTTCAAGATTTGGCTGGCCACCATAGTCT 360
 DB 384 TTCCCTTCTGGGAAGTGTGTATCACTGTTCAAGATTTGGCTGGCCACCATAGTCT 443
 QY 361 TTGTCGCTGATGATCTCAAGTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAG 420
 DB 444 ATGTCGCTGATGATCTCAAGTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAG 503
 QY 421 GTGATGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGCGGCTTCTAGCG 480
 DB 504 GTGATGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGCGGCTTCTAGCG 563
 QY 481 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 564 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 623
 QY 541 CTGTCGTGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGCGGCTTCTAG 600
 DB 624 CTGTCGTGCTGTGTGTCAGTGGGCAATCTGCTGAGCGGCTTCTAGCGGCTTCTAG 683
 QY 601 ACACGACGATCCCACTTTGTC 623
 DB 684 AGACGACGATCCCACTTTGTC 706

RESULT 7
 BX406634/c
 LOCUS BX406634 878 bp mRNA linear EST 03-MAY-2004
 DEFINITION BX406634 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CS0DM013YA17 3-PRIME, mRNA sequence.
 ACCESSION BX406634
 VERSION BX406634.2 GI:46951844
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 I.J.W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization

ORIGIN

sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
Http://www.nisc.nih.gov/).

Query Match 44.7%; Score 560.4; DB 5; Length 631;
Best Local Similarity 99.6%; Pred. No. 4.1e-143;
Matches 561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATAGAGCTTAAGTACCCGCGGTCTGTCGGGGGCTGCGCCCTCTGGGCGCTTAACATG 60
DB 69 ATGAGCTTAAGTACCCGCGGTCTGTCGGGGGCTGCGCCCTCTGGGCGCTTAACATG 128
QY 61 GAAGAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 129 GAAGAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 188
QY 121 CAAAAGGGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 180
DB 189 CAAAAGGGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 248
QY 181 GGGCTTGGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 240
DB 249 GGGCTTGGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 308
QY 241 CTCTTCATGCTGCGCTTGGTGTGTCGATGCGCAATCTCTGTCGACGCGCTTCTGAGCC 300
DB 309 CTCTTCATGCTGCGCTTGGTGTGTCGATGCGCAATCTCTGTCGACGCGCTTCTGAGCC 368
QY 301 TTCCCTTCTGGGAAGTGTGTCATCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
DB 369 TTCCCTTCTGGGAAGTGTGTCATCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 428
QY 361 TTGTCGGTCTGATCTCAGTGGATCTGTCTTTGGGGAAGTCACTTGGCGCAGTTGGT 420
DB 429 TTGTCGGTCTGATCTCAGTGGATCTGTCTTTGGGGAAGTCACTTGGCGCAGTTGGT 488
QY 421 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACTGAGAGATGTCATCAGTAATATC 480
DB 489 GTGATGGTGTGGTGGAGGTGACAGCTTTAGGCACTGAGAGATGTCATCAGTAATATC 548
QY 481 TTCAACACAGACTACCATGATGATGACATCTACCTGATGTCGAGCCTATTTTGG 540
DB 549 TTCAACACAGACTACCATGATGATGACATCTACCTGATGTCGAGCCTATTTTGG 608
QY 541 CTGTCGTGGCTGTGTCCTGCTGCC 563
DB 609 CTGTCGTGGCTGTGTCCTGCTGCC 631

```

RESULT 9
BU665450
LOCUS 632 bp mRNA linear EST 30-SHP-2002
DEFINITION BU665450.1 z1 Hembase; Erythroid Precursor Cells (LCB:cl library)
ACCESSION BU665450
VERSION BU665450.1 GI:23377637
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 632)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148

FEATURES

source

Email: jml7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 131 row: d column: 04
Seq primer: 5' lambda-triPLEX2 Sequencing Primer.
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C1131d04"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
library)"
/note="Organ: blood; Vector: pTriPLEX; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA library
construction kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/L peptide nucleic acid
(PNA) oligos
(N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
(N-terminal)-biotin-C(T/C)-T-GAA-GTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
Http://www.nisc.nih.gov/).

ORIGIN

Query Match 44.6%; Score 559; DB 5; Length 632;
Best Local Similarity 98.3%; Pred. No. 1e-142;
Matches 565; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 ATAGAGCTTAAGTACCCGCGGTCTGTCGGGGGCTGCGCCCTCTGGGCGCTTAACATG 60
DB 58 ATGAGCTTAAGTACCCGCGGTCTGTCGGGGGCTGCGCCCTCTGGGCGCTTAACATG 117
QY 61 GAAGAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 118 GAAGAGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 177
QY 121 CAAAAGGGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 180
DB 178 CAAAAGGGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 237
QY 181 GGGCTTGGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 240
DB 238 GGGCTTGGGCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 297
QY 241 CTCTTCATGCTGCGCTTGGTGTGTCGATGCGCAATCTCTGTCGACGCGCTTCTGAGCC 300
DB 298 CTCTTCATGCTGCGCTTGGTGTGTCGATGCGCAATCTCTGTCGACGCGCTTCTGAGCC 357
QY 301 TTCCCTTCTGGGAAGTGTGTCATCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
DB 358 TTCCCTTCTGGGAAGTGTGTCATCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 417
QY 361 TTGTCGGTGTGATCTCAGTGGATCTGTCTTTGGGGAAGTCACTTGGCGCAGTTGGT 420
DB 418 ACGTCGGTGTGATCTCAGTGGATCTGTCTTTGGGGAAGTCACTTGGCGCAGTTGGT 477

```


size selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; <http://www.nisc.nih.gov/>)."

Query Match	43.3%	Score 542.4	DB 5	Length 607
Best Local Similarity	98.0%	Pred. No. 3.8e-118		
Matches 549	Conservative 0	Mismatches 11	Indels 0	Gaps 0
Qy	1	ATGAGCTCTAAGTACCGCGGTCTGTCTCGCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG	60	
Db	48	ATGAGCTCTAAGTACCGGTGTGTCTCGCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG	107	
Qy	61	GAAGCAGCTCATTTCTCTCTTTATTTTTTTAACCACATATGACGCTTCTTATAGAGAT	120	
Db	108	GAAGCAGCTCATTTCTCTCTTTATTTTTTTAACCACATATGACGCTTCTTATAGAGAT	167	
Qy	121	CAAAAGGGGGCTCGTGGSCATCCTATCAAGTTGGCCAAAGTCTGACCGTATGGGGCAATT	180	
Db	168	CAAAAGGGGGCTCGTGGSCATCCTATCAAGTTGGCCAAAGTCTGACCGTATGGGGCAATT	227	
Qy	181	GGCTTTGGGCTTCTCAACCTCGAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC	240	
Db	228	GGCTTTGGGCTTCTCAACCTCGAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC	287	
Qy	241	CTTTTCATGTGCGCGCTTGTGTGTCAGTGGGCAATCTCTGTGACGCGCTTCTGAGCCAG	300	
Db	288	CTTTTCATGTGCGCGCTTGTGTGTCAGTGGGCAATCTCTGTGACGCGCTTCTGAGCCAG	347	
Qy	301	TTCCCTCTTGGGAAGGTGTATACACATGTTCAGATTTCCGCTGGCCACATAGATGCT	360	
Db	348	TTCCCTCTTGGGAAGGTGTATACACATGTTCAGATTTCCGCTGGCCACATAGATGCT	407	
Qy	361	TTGTGCGGTGATCTCAGTGGATGTCTTCTTGGGGAAAGTCAACTTGGGCGCAGTGGGT	420	
Db	408	ATGTGCGGTGATCTCAGTGGATGTCTTCTTGGGGAAAGTCAACTTGGGCGCAGTGGGT	467	
Qy	421	GTGATGTGTCTGTGGAGGTGACACGCTTTAGGCAACCTGAGGATGTCTATCAGTAATATC	480	
Db	468	GTGATGTGTCTGTGGAGGTGACACGCTTTAGGCAACCTGAGGATGTCTATCAGTAATATC	527	
Qy	481	TTCAACACACACTACACATATGATGACATCTACGTGTTCGACGCTATTTTGGG	540	
Db	528	TTCAACACACACTACACATATGATGACATCTACGTGTTCGACGCTATTTTGGG	587	
Qy	541	CTGTCTGTGGCGCTGTGCT	560	
Db	588	CTGTCTGTGGCGCTGTGCT	607	

RESULT 13

AK089642

LOCUS

DEFINITION

AK089642 975 bp mRNA linear HTC 03-APR-2004

Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone: F83006J05 product: Rhesus blood group CE and D, full insert sequence.

AK089642

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

3

REFERENCE

AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagakawa, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishii, K., Kitamura, T., Tashiro, H., Itoh, M., Samu, H., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Ozaki, Y., Murakatsu, M., Inoue, Y., Kira, A. and Hayaehizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	
TITLE	
JOURNAL	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length mouse cDNAs
AUTHORS	Nature 420, 563-573 (2002)
TITLE	6 (bases 1 to 975)
JOURNAL	
REFERENCE	
ADACHI, J., Aizawa, K., Akiyama, T., Arai, K., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hasegawa, W., Hayaehizaki, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Katsukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Onose, N., Ozaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaehizaki, Y.	
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayaehizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-research@riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	
SOURCE	1. .975
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="NOD"
	/db_xref="FANTOM DB:FB30006J05"
	/db_xref="taxon:10090"
	/clone_id="FB30006J05"
	/clone_type="activated spleen"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	19. .975
	/note="unnamed protein product; Rhesus blood group CE and D [MGD] MGI:1202882, GB NM_011270, evidence: BLASTN, 99%, match=1257
	putative"
	/codon_start=1
	/protein_id="BAC40937.1"


```

/db xref="GI:26354621"
/translation="MGSXPRLRCCLPLMALELOTAFSLFCFPIPHDTQVDRPM
ASYOVLRLTLMALGFLSSPRHSSVAFLNMLAGVOTIILDFLGVLQ
MKRINLSIOIATSLPLVLSAGAVLKVHVLDTVMLEAFALIFPADKVF
KQTEHLIYNHGFVAFYRLTYAMWLSLPRVENQTEKVNATSSLSLFLMIGTL
PLMTIPALNSHLSSTKRNAVENVYALAVSAVATSMALSHPOKINMVIHNA
VLAGVAVGAPGLISPWISWVLIIGLISWAGKPRACL"

```

ORIGIN

```

Query Match      39.8%; Score 498.6; DB 3; Length 975;
Best Local Similarity 72.4%; Pred. No. 5,4e-126;
Matches 696; Conservative 0; Mismatches 244; Indels 21; Gaps 3;

QY 1 ATGAGCTTAAGTACCCCGGCTGTGTCCGCGCTGCTCCCTCGGAGCCCTTAACATG 60
DB 19 ATGGGCTTAAGTACCCCGGCTGTGTCCGCGCTGCTCCCTCGGAGCCCTTAAGT 78
QY 61 GAAGAGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 79 CAGACAGCTTTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 138
QY 121 CAAAAGGGGCTGTGGCATCTCATCAAGTTGGCCAAAGTGTGACCGTATGGCGGCATT 180
DB 139 CACA---GGTTCATGGCAGGATCAAGTCTCCGGAATTTGACCTCATGACGCTTG 195
QY 181 GGCTTGGGCTTCTCACTCGAGTTCCGAGAACAGCTGAGCAAGTGTGGCTTGAAC 240
DB 196 GGCTTGGGCTTCTCACTCGAGTTCCGAGAACAGCTGAGCAAGTGTGGCTTGAAC 255
QY 241 CTCTTCATGTGGGCTGTGTGTGAGAGGCAATCCGTGTGAGCGGCTTCTTGAAGCAG 300
DB 256 CTCTTCATGTGGGCTGTGTGTGAGAGGCAATCCGTGTGAGCGGCTTCTTGAAGCAG 315
QY 301 TTCCCTTCTGGGAAAGGTGTGATCACTGTTCAGTATTCGGCTGACCATGAGTGT 360
DB 316 GTCTCCCAATGGAACAAGATCAACATCTGTCCAGCATCCAGATAGCTACCATGAGCACC 375
QY 361 TTGTGGGCTGTGATCTAGTGGATGCTGTGTGGGAAAGTCAACTTGGCGCAGTTGGTG 420
DB 376 TTACTGTGTGATCTAGTGGATGCTGTGTGGGAAAGTCAACTTGGCGCAGTTGGTG 435
QY 421 GTGATGTGTGTGAGAGTGAACGTTTAAAGCACTGAGGATGTGATCACTAATATTC 480
DB 436 GTGATGTGTGTGAGAGTGAACGTTTAAAGCACTGAGGATGTGATCACTAATATTC 495
QY 481 TTGCAACAACACTTCAACATGATGATGATCACTTATGCTGTGAGCTTATTTGGG 540
DB 496 TTCAAAATGACAAACAATCATCTGTGATGACGAGGACGTTGTTGGGGCTATTTGGG 555
QY 541 CTGTCTGTGGCTGTGCTGCAAAAGCTCTAACCAGAGGAGAGGATTAAGATCAG 600
DB 556 CTAACTGTGGCTGTGCTGCAAAAGCTCTAACCAGAGAGTGTGAGGAGAGGAGC 615
QY 601 ACA-----GCAACGATACCCAGTTTGTCTGCAATGCTGGGCGGCTCTTTC 645
DB 616 ACAAGAGAGTTCAAAATGCTAGAGCTCAAGCTGTGTTGCAATGCTGGGCAACCTCTTTC 675
QY 646 TTGAGATGTGTTGGGCAAGTTTCACTGCTGTCTGTGAAGTCAATTCGAAAGGAAAG 705
DB 676 TTGTGATATTTCTGGGCAAGTATCACTGCTCTCTGAAAGGGAACAAAGAAAG---G 732
QY 706 AATGCGGTGTTCAACCTACTATGCTGTAGCAGTCAAGGCTGTGACAGCCATTCAGGG 765
DB 733 AATGCTGTGTTCAACCTACTATGCTGTAGCAGTCAAGGCTGTGACAGCCATTCAGGG 792
QY 766 TCACTCTTGGCTCAACCCCAAGGAAAGTCAAGCACTTATGTGACAGTGTGGTGTG 825
DB 793 TCAAGCCCTAGTCAACCCCAAGGAAAGTCAACATGTTCAATCAACATGCAATGCTG 852
QY 826 GAGAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 885
DB 853 GAGAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 912

```

```

QY 886 GTGTGGGCTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 945
DB 913 GTGTGGGCTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 972
QY 946 T 946
DB 973 T 973

```

```

RESULT 14
CD627633/c 458 bp, mRNA linear EST 12-JUN-2004
LOCUS 56029924H1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD627633
ACCESSION CD627633
VERSION CD627633.1 GI:40275899
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 458)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
source 1..458
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_11b="FLP"
/note="Vector: pDrive Cloning Vector"

```

FEATURES

source

ORIGIN

```

Query Match      34.2%; Score 429.2; DB 6; Length 458;
Best Local Similarity 98.2%; Pred. No. 6.1e-107;
Matches 434; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCCGGCTGTGTCCGCGCTGCTCCCTCGGAGCCCTTAACATG 60
DB 442 ATGAGCTTAAGTACCCCGGCTGTGTCCGCGCTGCTCCCTCGGAGCCCTTAACATG 383
QY 61 GAAGAGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 382 GAAGAGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 323
QY 121 CAAAAGGGGCTGTGGGCAATCCTATCAAGTTGGGCAACATCTGACCGGATGAGGCGCAT 180
DB 322 CAAAAGGGGCTGTGGGCAATCCTATCAAGTTGGGCAACATCTGACCGGATGAGGCGCAT 263
QY 181 GGCTTGGGCTTCTCACTCGAGTTTCGGAAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
DB 262 GGCTTGGGCTTCTCACTCGAGTTTCGGAAGACACAGCTGAGAGAGTGTGGCTTCAAC 203
QY 241 CTCTTCATGTGGGCTGTGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 300
DB 202 CTCTTCATGTGGGCTGTGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 143
QY 301 TTCCCTTCTGGGAAAGTGTGATCACTGTGATTCAGTATTCGCTGTGCAACATGAGTGT 360
DB 142 TTCCCTTCTGGGAAAGTGTGATCACTGTGATTCAGTATTCGCTGTGCAACATGAGTGT 83
QY 361 TTGTGTGTGTGATTCAGTGTGATGTGCTGTGTGGGAAAGTCAACTTGGCGGAGTTGGTG 420
DB 82 TTGTGTGTGTGATTCAGTGTGATGTGCTGTGTGGGAAAGTCAACTTGGCGGAGTTGGTG 23
QY 421 GTGATGTGTGTGATGTGAGAGTGA 442

```


Db 22 GTGAGGTCAACAGGTGAGTCA 1

RESULT 15
BU663660 612 bp mRNA linear EST 30-SEP-2002
LOCUS c1104e12.21 Hembase; Erythroid Precursor Cells (LCB:c1 library)
DEFINITION Homo sapiens cDNA clone c1104e12 5', mRNA sequence.
ACCESSION BU663660
VERSION BU663660.1 GI:23375845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 612)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jlm7@nih.gov
The 'c1' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 104 row: e column: 12
Seq primer: 5' lambda-Triplex2 Sequencing primer.
Location/Qualifiers
1. 612
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c1104e12"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA+"
/lab_host="DH5alpha"
/clone_id="Hembase; Erythroid Precursor Cells (LCB:c1
library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA library
Construction Kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/l peptide nucleic acid
(PNA) oligos
(N-terminal) -biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and
(N-terminal) -biotin-C(T/C)T-GAA-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
Http://www.nisc.nih.gov/)."

ORIGIN

Query Match 34.1%; Score 427.2; DB 5; Length 612;
Best Local Similarity 89.6%; Pred. No. 2,4e-106;
Matches 489; Conservative 0; Mismatches 13; Indels 44; Gaps 1;

QY 6 CTCTAAGTACCCGGGCTGTCTCGCGCTGCTGCCCCCTTGAGCCCTTAACACTGGAAGC 65
|||||

Db 67 CTCTAAGTACCCGGGCTGTCTCGCGCTGCTGCCCCCTTGAGCCCTTAACACTGGAAGC 126
QY 66 AGCTCATTTCTCTCTCTTCTATTTTAAACCACTAGACGCTTCTTAGAGATCAAAA 125
Db 127 AGCTCATTTCTCTCTCTTCTATTTTAAACCACTAGAGCTTCTTAGAGATCAAAA 186
QY 126 GGGGCTGTGGCATCTTATC----- 145
Db 187 GGGGCTGTGGCATCTTATCAGAGGGGCTGGGAAGTGTGATGCTGTATAAATCCAGAA 246
QY 146 ----AAGTGGCCAAATCTGACCGTATGAGCGGCTTGGCTTCCCTCACTCG 201
Db 247 CCAGAACTGGCCAAAGATCTGACCGTATGAGCGGCTTGGCTTCCCTCACTCA 306
QY 202 AGTTTCGGAGACACAGCTGAGACAGTGGCTTCACTTTCATGCTGGCGCTTGGT 261
Db 307 AATTTCCGAGACACAGCTGAGACAGTGGCTTCACTTTCATGCTGGCGCTTGGT 366
QY 262 GTGCAAGTGGCAATCTCTGTCGACGCGCTTCTGAGCCAGTTCCCTTGGGAAAGTGTG 321
Db 367 GTGCAAGTGGCAATCTCTGTCGACGCGCTTCTGAGCCAGTTCCCTTGGGAAAGTGTG 426
QY 322 ATCACAAGTTCAGTATTCGCTGGCCAGCATGAGTCTTTCGCTGATTCAGTG 381
Db 427 ATCACAAGTTCAGTATTCGCTGGCCAGCATGAGTCTTTCGCTGATTCAGTG 486
QY 382 GATGCTGCTTGGGAAAGTCACTTGGCGCAGTGGTGTGATGGTGTGAGAGTG 441
Db 487 GTGCTGCTTGGGAAAGTCACTTGGCGCAGTGGTGTGATGGTGTGAGAGTG 546
QY 442 ACAGCTTTAGCAACCTGAGATGTCATGATTAATCTTCAACACAGACTACACATG 501
Db 547 ACAGCTTTAGCAACCTGAGATGTCATGATTAATCTTCAACACAGACTACACATG 606
QY 502 AACATG 507
Db 607 AACCTG 612

Search completed: March 23, 2005, 03:45:41
Job time : 3950 secs

This Page Blank (uspto)